



A University of Sussex DPhil thesis

Available online via Sussex Research Online:

<http://sro.sussex.ac.uk/>

This thesis is protected by copyright which belongs to the author.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Please visit Sussex Research Online for more information and further details

Psychophysiological Indices of Recognition Memory

Becky Heaver

Thesis submitted for the degree of Doctor of Philosophy

University of Sussex, December 2011

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature:

Date:

Acknowledgements

I owe my deepest gratitude to my DPhil supervisor, Dr Sam Hutton, without whom this thesis simply would not have been possible, and who has made available his support in a number of ways, and encouraged, guided and supported me from green masters student to completed doctoral candidate. I would also like to thank him for, amongst other things, his hardcore binary re-programming of a parallel cable, a number of interesting side projects, and never giving me cause to attend a 'How to Manage Your Supervisor' workshop. I would like to thank Francesca Citron, Kathrin Mikan and Steve Hamilton for their tuition and assistance in ERP techniques. Special thanks go to Jacob Readman, Emily Lewis and Rakshita Roplekar, University of Sussex, for helping with data collection during Experiments 1, 3, 6 and 7, and to all my participants, many of whom were friends who helped out of the goodness of their hearts, and some of whom managed to stay awake for the duration of entire experiments. I also thank Prof Alan Garnham for asking me questions that really made me think; Prof Brennis Lucero-Wagoner who truly touched me by looking through 'the detritus of 40 years of academic life' just to find me a technical report in my first year when she didn't know me; Prof Stuart Steinhauer for being so approachable when I was star-struck; everybody at the PIL lab, Stirling University, for giving me confidence in my ERP data, and pointing me in the right direction to develop my skills further; Pennie Ingram for everything she does behind the scenes; Prof Angie Hart for working so hard to understand me, accommodate me and help me grow; Sarah Hendrickx and Victoria Franklin for showing me who I am at the

point when I probably needed it the most; Terry Buckley for his assistance with my Latin references. I am indebted to the friends, family and colleagues who have supported me, especially Anne and Malcolm, Trish and Mitch, Ben, Riley and Toby, Sue, Anneka, Nikki, Clare, Gary, Rebecca, Munira, JJQ, Graham, Ringo, Sam C, Sarah RH, Martha, Toni, Chrissy, Kim, Dan, Alice, Jo, Prozac, alcohol... and of course my cats Dizzy, Smoo and Lennie for their constant unconditional love, support and amusement. Because it has been essential to maintaining my motivation, I would like to thank 'past me' for making copious notes whilst reading things I knew 'future me' was going to need whilst writing up. My ultimate thanks goes to Alex for going to the shop to get wine more than his fair share of the time, taking over the daily cooking and chores, putting up with me no matter what, and giving me space whilst still being close. Without these people I could not have achieved what I have, and I offer my regards and thanks to everybody else who supported me in any respect during the completion of my thesis.

Publications

The following journal articles and conference presentations have been adapted from experimental work detailed in this thesis:

- Heaver, B., & Hutton, S.B. (2011). Keeping an eye on the truth? Pupil-size changes associated with recognition memory. *Memory*, 19(4), 398-405.
doi:10.1080/09658211.2011.575788
- Heaver, B., & Hutton, S.B. (2010). Keeping an eye on the truth: Pupil-size, recognition memory and malingering. *International Journal of Psychophysiology*, 77(3), 306. doi:10.1016/j.ijpsycho.2010.06.206
- Heaver, B., & Hutton, S.B. (2010). Keeping an eye on the truth: Pupil-size, recognition memory and malingering. Poster presented at The 15th World Congress of Psychophysiology of the International Organization of Psychophysiology (I.O.P.), Budapest, Hungary, 1-4 September, 2010.
- Heaver, B., & Hutton, S.B. (2009). Keeping an eye on the truth. Oral presentation at the Annual Conference of the British Psychological Society, Brighton, UK, 1-3 April, 2009.
- Heaver, B., & Hutton, S.B. (2008). You can't hide your lying eyes: Pupillary responses during recognition memory. Poster presented at the Cognitive Section Conference of the British Psychological Society, Southampton, UK, 9 September, 2008.

Thesis Summary

UNIVERSITY OF SUSSEX

BECKY HEAVER

THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

PSYCHOPHYSIOLOGICAL INDICES OF RECOGNITION MEMORY

SUMMARY

It has recently been found that during recognition memory tests participants' pupils dilate more when they view old items compared to novel items. This thesis sought to replicate this novel "Pupil Old/New Effect" (PONE) and to determine its relationship to implicit and explicit mnemonic processes, the veracity of participants' responses, and the analogous Event-Related Potential (ERP) old/new effect. Across 9 experiments, pupil-size was measured with a video-based eye-tracker during a variety of recognition tasks, and, in the case of Experiment 8, with concurrent Electroencephalography (EEG). The main findings of this thesis are that:

- the PONE occurs in a standard explicit test of recognition memory but not in "implicit" tests of either perceptual fluency or artificial grammar learning;
- the PONE is present even when participants are asked to give false behavioural answers in a malingering task, or are asked not to respond at all;

- the PONE is present when attention is divided both at learning and during recognition;
- the PONE is accompanied by a posterior ERP old/new effect;
- the PONE does not occur when participants are asked to read previously encountered words without making a recognition decision;
- the PONE does not occur if participants preload an “old/new” response;
- the PONE is not enhanced by repetition during learning.

These findings are discussed in the context of current models of recognition memory and other psychophysiological indices of mnemonic processes. It is argued that together these findings suggest that the increase in pupil-size which occurs when participants encounter previously studied items is not under conscious control and may reflect primarily recollective processes associated with recognition memory.

Contents

Acknowledgements	ii
Publications	iv
Thesis Summary	v
List of Figures	xi
1. Introduction – Pupil-Size and Cognitive Function	1
1.1. The Pupil	5
1.1.1. Anatomy & Physiology	5
1.1.1.1. Pharmacology	7
1.1.1.2. Pathology	9
1.1.2. Pupil-Size Change	10
1.1.2.1. The Light Reflex	12
1.1.2.2. The Darkness Reflex	13
1.1.2.3. The Lid-Closure Reflex	14
1.1.2.4. The Accommodation Response	14
1.1.2.5. Pupillary Hippius	15
1.1.2.6. Iris Colour	16
1.1.2.7. Role of the Locus Coeruleus in Stimulus-Evoked Dilations	17
1.2. Pupillometry Research Literature	20
1.2.1. Eckhard H. Hess	22
1.2.1.1. Criticisms of Hess and Early Work	24
1.2.2. Arousal	28
1.2.2.1. Fatigue & Sleepiness	32
1.2.3. “Cognitive Effort” – Kahneman	33
1.2.3.1. Signal Detection & Discrimination	39
1.2.3.2. Working Memory	40
1.2.3.3. Visual Search	41
1.2.3.4. Effort and the Red Pupillary Reflex	42
1.2.3.5. Language & Comprehension	42
1.2.3.6. Imagery	43
1.2.3.7. Lie Detection	43
1.2.3.8. Auditory Stimuli	44
1.2.4. Attention	44
1.2.4.1. Locus Coeruleus	46
1.2.4.2. Blinks	47
1.2.4.3. Decision Making & Uncertainty	48
1.2.5. Physical Effort	48
1.2.6. Pupil-Size and Concurrent Psychophysiological Measures	50
1.2.7. Memory	54
1.3. Recognition Memory	54
1.3.1. Models of Recognition Memory	54
1.3.1.1. Single-Process Models	57
1.3.1.2. Dual-Process Models	62
1.3.1.3. Evaluating Models	66
1.3.2. Psychophysiological Correlates of Recognition Memory Processes	70
1.3.2.1. Event-Related Potentials	71
1.3.2.2. PET and fMRI	72
1.3.2.3. Pupil-Size	72
1.4. Summary	79
2. Methods	82
2.1. Pupillometry	82
2.1.1. Background	82

2.1.1.1.	Techniques	82
2.1.1.2.	Data Acquisition	85
2.1.2.	Pupil-Size Reporting Variables	86
2.1.2.1.	Pupil Dilation Ratio	89
2.1.2.2.	Measurement Issues	91
2.1.3.	Pupil-Size Analysis	94
2.2.	Event-Related Potentials (ERPs)	97
2.2.1.	ERP Data Acquisition	97
2.2.2.	ERP Data Analysis	101
2.3.	Stimuli and Participants	103
2.3.1.	Word Selection	103
2.3.2.	Selection of Participants	104
3.	Replicating and Extending the Pupil Old/New Effect	106
3.1.	Experiment 1 – Implicit vs. Explicit Tests of Recognition	111
3.1.1.	Method	111
3.1.1.1.	Participants	111
3.1.1.2.	Materials/Apparatus	111
3.1.1.3.	Design and Procedure	112
3.1.1.4.	Pupil Recording	113
3.1.2.	Results	114
3.1.2.1.	Behavioural Data: Old/New Responses	114
3.1.2.2.	Behavioural Data: Confidence	116
3.1.2.3.	Pupil-Size Data	117
3.1.2.4.	Pupil-Size Data: Confidence Analysis	119
3.1.3.	Discussion	120
3.2.	Experiment 1b – Reading Condition	121
3.2.1.	Method	121
3.2.1.1.	Participants	121
3.2.1.2.	Materials/Apparatus	121
3.2.1.3.	Design and Procedure	121
3.2.1.4.	Pupil Recording	122
3.2.2.	Results	122
3.2.2.1.	Behavioural Data	122
3.2.2.2.	Pupil-Size Data	122
3.2.3.	Discussion	123
3.3.	Experiment 2 – Short vs. Long Presentation Duration	126
3.3.1.	Method	128
3.3.1.1.	Participants	128
3.3.1.2.	Materials/Apparatus	128
3.3.1.3.	Design and Procedure	129
3.3.1.4.	Pupil Recording	131
3.3.2.	Results	131
3.3.2.1.	Behavioural Data	131
3.3.2.2.	Pupil-Size Data	133
3.3.3.	Discussion	136
3.4.	General Discussion	137
4.	The Role of Conscious Awareness in the PONE	142
4.1.	Experiment 3 – Implicit Grammar	148
4.1.1.	Method	148
4.1.1.1.	Participants	148
4.1.1.2.	Materials/Apparatus	148
4.1.1.3.	Design and Procedure	150
4.1.1.4.	Pupil Recording	152
4.1.2.	Results	152
4.1.2.1.	Behavioural Data	152
4.1.2.2.	Pupil-Size Data	152
4.1.3.	Discussion	153
4.2.	Experiment 4 – Implicit Grammar Replication	154
4.2.1.	Method	154
4.2.1.1.	Participants	154

4.2.1.2.	Materials/Apparatus/Design/Procedure	154
4.2.2.	Results	155
4.2.2.1.	Behavioural Data	155
4.2.2.2.	Pupil-Size Data	156
4.3.	General Discussion	158
5.	Malingering and the Old/New Response	161
5.1.	Malingering and Deception	161
5.1.1.	Background	161
5.1.1.1.	Malingered Memory-Impairment	163
5.1.1.2.	Psychophysiological Detection of Deception	167
5.1.1.3.	Pupillometry Studies	174
5.2.	Experiment 5 – Malingering and the PONE	180
5.2.1.	Method	180
5.2.1.1.	Participants	180
5.2.1.2.	Materials/Apparatus	181
5.2.1.3.	Design and Procedure	181
5.2.1.4.	Pupil Recording	183
5.2.2.	Results	183
5.2.2.1.	Behavioural Data: Old/New Responses	183
5.2.2.2.	Behavioural Data: Confidence	184
5.2.2.3.	Pupil-Size Data	185
5.2.2.4.	Pupil-Size Data: Confidence Analysis	187
5.2.3.	Discussion	188
5.3.	Experiment 6 – Methods of Malingering	192
5.3.1.	Method	193
5.3.1.1.	Participants	193
5.3.1.2.	Materials/Apparatus	194
5.3.1.3.	Design and Procedure	194
5.3.1.4.	Pupil Recording	196
5.3.2.	Results	196
5.3.2.1.	Behavioural Data	196
5.3.2.2.	Pupil-Size Data	197
5.3.3.	Discussion	198
5.4.	Experiment 7 – Emulating Memory-impairment	202
5.4.1.	Method	203
5.4.1.1.	Participants	203
5.4.1.2.	Materials/Apparatus	204
5.4.1.3.	Design and Procedure	205
5.4.1.4.	Pupil Recording	206
5.4.2.	Results	206
5.4.2.1.	Behavioural Data	206
5.4.2.2.	Pupil-Size Data	208
5.4.3.	Discussion	209
5.4.4.	General Discussion	209
6.	ERP and Pupil Old/New Effects	214
6.1.	Introduction	214
6.1.1.	Background to ERPs and Recognition Memory	214
6.1.2.	Pupil Responses and ERPs	221
6.2.	Experiment 8 – Strength of Memory Effect	226
6.2.1.	Method	227
6.2.1.1.	Participants	227
6.2.1.2.	Materials/Apparatus	227
6.2.1.3.	Design and Procedure	228
6.2.1.4.	Pupil Recording	230
6.2.1.5.	Electrophysiological Recording and Analysis	230
6.2.2.	Results	231
6.2.2.1.	Behavioural Data	231
6.2.2.2.	Pupil-Size Data	232
6.2.2.3.	Event-Related Potentials	233
6.2.2.4.	Old/New Effects for Weak and Strong Items	236

6.2.2.5.	Weak Old vs. New Items	237
6.2.2.6.	Strong Old vs. New Items	239
6.2.2.7.	Effect of Presentation Frequency on Old/New Effect	241
6.2.2.8.	Early Effects (80-150ms)	242
6.2.3.	Discussion	243
6.3.	Experiment 9 – Pupil and Behavioural Data Only	245
6.3.1.	Method	245
6.3.1.1.	Participants	245
6.3.1.2.	Materials/Apparatus/Design/Procedure	246
6.3.1.3.	Pupil Recording	246
6.3.2.	Results	246
6.3.2.1.	Behavioural Data	246
6.3.2.2.	Pupil-Size Data	247
6.3.3.	Discussion	248
6.3.4.	General Discussion	249
7.	General Discussion	258
7.1.1.	Summary	258
7.1.2.	Future Directions	268
8.	Bibliography	272
9.	Appendices	341

List of Figures

Figure 1-1: Excerpt from Bartholomew's <i>De proprietatibus rerum</i> (1240/1483).....	1
Figure 1-2: Excerpt from Harvey's <i>Exercitatio duae anatomica de circulatione sanguinis</i> (1649) 4	
Figure 1-3: Two muscle groups regulate the size of the pupil.	6
Figure 1-4: Pupillary constriction and dilation pathways.	8
Figure 1-5: The accommodation reflex – the pupil constricts as the lens thickens.....	15
Figure 1-6: Concurrent pupil-size and LC neuron recording in the monkey.....	18
Figure 1-7: Percentage pupil-size change in relation to baseline pupil diameter.	23
Figure 1-8: Mean pupil-size in relation to 10 second lifting and resting periods.....	49
Figure 1-9: Human memory systems.	55
Figure 1-10: Decision space for the remember-know task.....	66
Figure 2-1: (a) Head-mounted and (b) tower-mounted EyeLink eye-trackers.....	84
Figure 2-2: Pupil as located by eye-tracking software	85
Figure 2-3: Experiment reported as (a) absolute pupil-size values (b) relative values.	89
Figure 2-4: Sources of variation in measurements of pupil diameter.....	92
Figure 2-5: Traditional 10-20 and 10-10 electrode configurations.	97
Figure 2-6: Geodesic sensor net 64 and 128 channel electrode maps.	98
Figure 2-7: 128 channel Geodesic sensor net worn by models.	98
Figure 2-8: Modified combinational nomenclature for the 10-10 system	99
Figure 2-9: The 10-10 system overlaid on the Geodesic 128 sensor net.	100
Figure 3-1: Proportion of correct responses in each condition	114
Figure 3-2: Proportion of correct responses to surface and semantic items in each condition	115
Figure 3-3: Average confidence rating for old and new responses to old and new items in the explicit condition.....	116
Figure 3-4: Pupil dilation ratio for old and new items in each condition	117
Figure 3-5: Pupil dilation ratio for new, surface and semantic items in each condition.....	119
Figure 3-6: Proportion of correct responses in each condition	122
Figure 3-7: Pupil dilation ratio for old and new items in each condition.	123
Figure 3-8: Proportion of correct responses for old and new items in each condition..	132
Figure 3-9: Pupil dilation ratio for old and new items in all four conditions	134
Figure 3-10: Pupil dilation ratio for correctly identified old and new items in all four conditions	136
Figure 4-1: The two finite state grammars used to generate the strings.....	144
Figure 4-2: Pupil dilation ratio for old and new stimuli in both conditions.....	153
Figure 4-3: Proportion of correct responses.	156
Figure 4-4: Pupil dilation ratio to old and new items in each condition	157
Figure 4-5: Pupil dilation ratio to correctly identified old and new items in each condition.	158

<i>Figure 5-1: Proportion of old responses to old and new items for standard and malingering conditions.....</i>	<i>184</i>
<i>Figure 5-2: Average confidence rating for correct old and new items in each condition.....</i>	<i>185</i>
<i>Figure 5-3: Pupil dilation ratio for old and new items in standard, malingering and single response conditions.....</i>	<i>186</i>
<i>Figure 5-4: Proportion of correct responses to old and new items for Standard, Incomplete Effort and Random conditions.....</i>	<i>197</i>
<i>Figure 5-5: Pupil dilation ratio for old and new items in Standard, Incomplete Effort, Random and Quiet conditions</i>	<i>198</i>
<i>Figure 5-6: Proportion of correct responses to old and new items for all conditions.</i>	<i>207</i>
<i>Figure 5-7: Proportion of correct responses to all items for the four combinations of single and learning task conditions.....</i>	<i>207</i>
<i>Figure 5-8: Pupil dilation ratio for old and new items in all conditions</i>	<i>208</i>
<i>Figure 6-1: Experimental procedure.....</i>	<i>229</i>
<i>Figure 6-2: Proportion of correct responses to new, weak and strong items</i>	<i>232</i>
<i>Figure 6-3: Pupil dilation ratio for correctly identified new, weak and strong items.</i>	<i>233</i>
<i>Figure 6-4: Grand average ERPs for correctly identified new, weak (presented once) and strong (presented three times) items at midline frontal, central and parietal electrodes.</i>	<i>234</i>
<i>Figure 6-5: Topographical distribution of old/new differences in mean amplitude (μV) for weak items (first row) and strong items (second row) between 300-700ms.</i>	<i>235</i>
<i>Figure 6-6: Mean amplitude for new and weak items at frontal, central and parietal electrodes at 300-500ms.</i>	<i>237</i>
<i>Figure 6-7: Mean amplitude for new and weak items at frontal, central and parietal electrodes at 500-700ms.</i>	<i>238</i>
<i>Figure 6-8: Mean amplitude for new and strong items at frontal, central and parietal electrodes at 300-500ms.....</i>	<i>239</i>
<i>Figure 6-9: Mean amplitude for new and strong items at frontal, central and parietal electrodes at 500-700ms.</i>	<i>240</i>
<i>Figure 6-10: Proportion of correct responses to new, weak and strong items</i>	<i>247</i>
<i>Figure 6-11: Pupil dilation ratio for correctly identified new, weak and strong items.</i>	<i>248</i>

1. Introduction – Pupil-Size and Cognitive Function

“...the eye is not only a passive organ and one of the gateways of knowledge, but is also a portal through which the working of the brain becomes manifest.” Samuel Wilks, 1883, p. 5-6.

The relationship between the eye and the inner workings of the mind has long been the subject of philosophy and art, conjuring up powerful imagery and inspiring poems, songs, and novels. Centuries before the causes of pupil-size changes were contemplated in scientific and other literature, Franciscan monk Bartholomew The Englishman wrote about the origins of the word “pupil” in his encyclopaedia *On the Properties of Things*, thought to have been written in the 1240s (Keen, 2007; see Figure 1-1):

pupilla secundum Isido. li. xij. c. ij.
est mediū punctus oculi ubi est vis
videndi in qua quia parue imagi-
nes videntur pupilla appellat. parvuli enim
pulli appellant. Vocatur autem pupilla eo q
pura sit et impolluta ut puella.

Figure 1-1: Excerpt from Bartholomew's *De proprietatibus rerum* (1240/1483).

Janisse cites a translation by Trivisa (1495), which says, “the blacke of theye ... is callyd Pupilla in latyn for small, ymages ben seen therin” (p. 1). Whilst Janisse (1977) described this as one of the earliest references to the pupil, in the original text Bartholomew refers to the much earlier etymology of seventh century Archbishop and historian Isidore of Seville (560-636). Isidore believed the eyes were the sensory

organ closest to the soul: "... every indication of the mental state is in the eyes, whence both distress and happiness show in the eyes" (Isidore, 636/2006, XI.i.36).

The section of Isidore's etymology that Bartholomew refers to reads:

"The pupil (pupilla) is the middle point of the eye in which the power of vision resides; because small images appear to us there, they are called pupils, since small children are called pupils. There are many who use the form pupula, but it is called pupilla because it is pure and unpolluted, just like 'young girls'." (Isidore, 636/2006, XI.i.37).

Derived from the Latin word "pupilla", meaning "young girls", the pupil is named for the tiny reflections of people that can be seen in someone's eyes. In fact Isidore based his works on the writings of earlier scholars, such as the Roman naturalist and historian Gaius Plinius Secundus (Pliny the Elder, 23-79 AD) who wrote in his *Natural History* encyclopaedia: "...the small pupil can reflect the entire image of a human being" (Pliny, 79/1938, XI.LV). Pliny also commented on the mind-eye connection, although he too did not explicitly link it to the pupil:

"Nobody has eyes of only one colour ... No other part of the body supplies greater indications of the mind – this is so with all animals alike, but specially with man – that is, indications of self-restraint, mercy, pity, hatred, love, sorrow, joy." (Pliny, 79/1938, XI.LIV).

From the earliest writings, these sentiments have been echoed by poets, philosophers and authors including Guillaume de Salluste (1544-1590), Joshua Sylvester (1563-1618), Shakespeare (1564-1616), Descartes (1596-1650), Johann Kaspar Lavater (1741-1801), Byron (1788-1824) and Tennyson (1809-1892), all of whom have written about the possibility of our eyes revealing the state of our mind (Andreassi, 2000; Clark, 1885; Loewenfeld, 1958; Wilks, 1885). From the sixteenth century onward there is a marked increase in the amount of writing about the pupil

(Loewenfeld, 1958). Berrien and Huntington (1943) argue that the pupil had been seen as an emotional barometer since at least the 1800s, and that it had long been associated with “arousal”. Honoré de Balzac (1799-1850) and Charlotte Brontë (1816-1855) appear to be the first to specifically mention changes in pupil-size in their texts (Clark, 1885; Wilks, 1883; 1885). In *Jane Eyre* Brontë (1847/1946) wrote, "Pain, shame, ire, impatience, disgust, detestation, seemed momentarily to hold a quivering conflict in the large pupil dilating under his ebon eyebrow" (p. 163). De Balzac (1841/2000) described pupil constriction in response to positive emotions, “The blue of the iris expanded like a flower, diminishing the dark circle of the pupil, and seeming to float in a liquid and languishing light that was full of love” (p. 31). He also links dilation of the pupil with negative emotions:

When Monsieur de Grandville... whom she declined to take as a husband, kissed her hand with an earnest expression of regret, the new bishop noticed the strange manner in which the black pupil of Veronique's eyes suddenly spread over the blue of the iris, reducing it to a narrow circle. The eye betrayed unmistakably some violent inward emotion. (de Balzac, 1841/2000, pp. 79-80).

Additionally, he asks whether dilation might reflect passion:

The pupils of her eyes, gifted with the power of great expansion, widened until they covered the whole surface of the blue iris except for a tiny circle... Was it the storm of restrained passions; was it some power coming from the depths of the soul, which enlarged the pupils in full daylight as they sometimes in other eyes enlarge by night, darkening the azure of those celestial orbs?” (de Balzac, 1841/2000, p. 11).

In a thought piece to *Brain* in 1883, and again in a letter to *Nature* in 1885, eminent physician to Queen Victoria, Sir Samuel Wilks, drew attention to converging sources of evidence from doctors and physiognomists, and his own observations of the mood

of his pet parrot, that in addition to responding to emotions, the pupil may manipulate them in others (Wilks, 1883; 1885). For example, Foster wrote in his *Textbook of Physiology* (1891) that the pupil dilates “as an effect of emotions” (p. 1172).

Whilst much of the early pupil literature comprises either medical accounts of the physiological changes, or literary descriptions of the pupil-size changes accompanying their characters’ emotions, the earliest scientific report of pupil changes in response to internal states referred to by Wilks (1885) was by another physician, William Harvey (1578-1657), who wrote the first complete and detailed account of the circulatory system in the seventeenth century. Harvey writes, “In anger the eyes are fiery, and the pupils contracted” (Harvey, 1649, p. 152; see Figure 1-2). Fontana (1765, cited in Loewenfeld, 1958) is thought to have provided the first detailed account of the psychological stimuli that dilate the pupil, referred to in older texts as dilatation.

Ira rubent oculi, constringitur pupilla

Figure 1-2: Excerpt from Harvey’s *Exercitatio duae anatomica de circulatione sanguinis* (1649).

Parallel advances in the understanding of basic neurophysiology meant that the relationship between pupil-size and the autonomic nervous system was established by the 1850s (e.g., Bernard, 1852; Budge, & Waller, 1851; Kuntz, 1929). Although Wilks (1885) felt that dilated pupils were associated with relaxed contemplation, and constricted pupils with concentration, Fontana (1765, cited in Loewenfeld, 1958), Gratiolet (Gratiolet, & Grandeau, 1865) and Hack Tuke (1884) proposed that fear caused dilation, while Clark (1885) posited that any “strong mental emotion” would produce dilation of the pupil (p. 433). The discussion even drew the attention of

Charles Darwin in *Expression of the Emotions in Man and Animals* (1872), although he remained unconvinced and asked that further research be conducted.

Thus, even from the earliest writings, there appears to have been some agreement that there are numerous non-luminance-based influences on pupil-size, but considerable debate as to exactly what those influences are, and what their precise effects on pupil-size are – a situation that some may argue remains today. This thesis examines a relatively novel psychophysiological index of recognition memory, the Pupil Old/New Effect (PONE). This chapter will provide the theoretical background for the experiments presented in Chapters 3-6, starting with a brief description of the anatomy and physiology of the pupil, followed by a summary of the history of pupillometry, and research into the effects of cognitive processes on pupil-size up to the present day. The next section introduces some current models of recognition memory and describes some of the key paradigms researchers have used to test them. This is followed by a consideration of other psychophysiological correlates of recognition memory, such as the Event-Related Potential (ERP) old/new effect, which sets the scene for a comprehensive exploration of the literature on the PONE to date.

1.1. The Pupil

1.1.1. Anatomy & Physiology

The pupil is the circular aperture at the centre of the iris which allows light from our environment to pass freely to the light-sensitive sensory cells of the retina. The size of the pupil is determined by the iris muscles, the radial dilator pupillae and the concentric smooth muscle circles of the sphincter pupillae, which work in opposition

to dilate (mydriasis) and constrict (miosis) the pupil (see Figure 1-3; Löwenstein, & Loewenfeld, 1962; Beatty, & Lucero-Wagoner, 2000; Reeves, & Swenson, 2004).

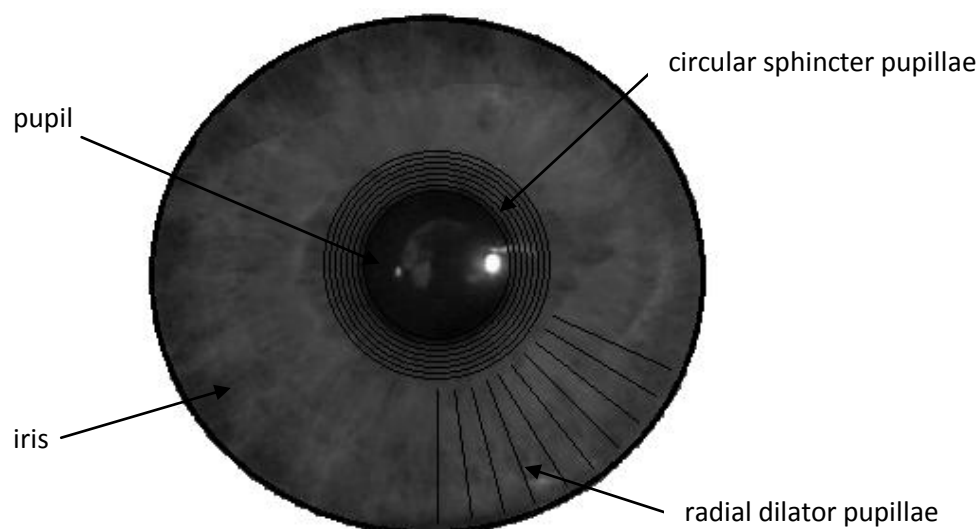


Figure 1-3: Two muscle groups regulate the size of the pupil: circular sphincter pupillae contract to make the pupil smaller, radial dilator pupillae contract to make the pupil larger.

The sphincter and dilator muscles are innervated by the Autonomic Nervous System (ANS) parasympathetic and sympathetic pupillomotor fibres of the third cranial nerve respectively (Foster, 1891; Reeves, & Swenson, 2004). The sympathetic and parasympathetic systems work in opposition, and tonic activation in both systems balance each other to produce an average waking pupil-size in ambient illumination of 2-6mm, with an average of 5mm and a range of 1-9mm (Beatty, & Lucero-Wagoner, 2000; Reeves, & Swenson, 2004). An increase in efferent activity to either muscle group leads to increasing central inhibition of the motor-nucleus of the other group (Miller, & Newman, 2005).

It has been shown that average 'resting' pupil-size decreases curvilinearly from about age twenty, becoming asymptotic at around sixty years (Birren, Casperson, & Botwinick, 1950), and that pupil-size becomes more variable with age (Kumnick, 1954; 1956). It has been suggested that this may be due to increased rigidity of the

sphincter muscle (Reeves, & Swenson, 2004; Spector, 1990) or decreasing sympathetic activity (Bitsios, Prettyman, & Szabadi, 1996a). The pupils are ‘yoked’, so that a change in one pupil shows a consensual response in the other because each retina inputs equally to the pretectal region and Edinger-Westphal nuclei, and the sphincter muscle of each pupil receives equal efferent output from these brain regions (Reeves, & Swenson, 2004). Around one person in four has naturally-occurring benign asymmetry of the pupils of up to 0.5mm, known as anisocoria (Reeves, & Swenson, 2004).

1.1.1.1. *Pharmacology*

The effects of certain substances on the pupil have been known since Ancient Roman times, when the juices of the poisonous plant *Atropa belladonna* (deadly nightshade) were used to dilate the pupil for cataract surgery (Loewenfeld, 1958). Apparently lost, this knowledge was rediscovered in the seventeenth century when ladies purportedly applied the plant to their eyes so that their dilated pupils would make them appear more attractive to their unsuspecting admirers (belladonna meaning ‘beautiful woman’ in Italian) (Forbes, 1977; Wilks, 1883; Wootton, 1910). Belladonna contains atropine, which blocks the parasympathetic input to the sphincter muscle.

Parasympathetic nerves transmit messages relating to “rest and digest” functions, using Acetylcholine (ACh) neurotransmitter both centrally and peripherally, and the efferent fibres innervating the sphincter muscle of the iris originate in the midbrain Edinger-Westphal nucleus (Figure 1-4). ACh acts on the sphincter via muscarinic receptors, leading to constriction of the pupil (Fountoulakis, 1999). Topically applied pharmacological ACh agonists, such as pilocarpine, and ACh-breakdown inhibitors, such as physostigmine, cause constriction of the sphincter, whereas ACh antagonists, such as atropine, relax the sphincter, enhancing the efforts of the dilator

muscle (Miller, & Newman, 2005). Miosis can also be induced centrally by the effects of chloroform, sedation, and through direct stimulation of the Edinger-Westphal nucleus by morphine and other opioids (Foster, 1891; Knaggs, Crighton, Cobby, Fletcher, & Hobbs, 2004; Miller, & Newman, 2005).

Sympathetic fibres carry signals relating to “fight or flight” responses, using ACh centrally and Norepinephrine (NE, also known as noradrenaline) peripherally, and the efferent fibres innervating the dilator muscle originate in the hypothalamic motor area of the diencephalon (Figure 1-4). NE acts on the dilator via α -adrenergic receptors, leading to dilation of the pupil (Fountoulakis, 1999). NE agonists, such as ephedrine, and NE-reuptake inhibitors such as cocaine, cause dilation of the pupil, whereas NE antagonists, such as thymoxamine, relax the dilator muscle and reduce its opposition of the sphincter. Mydriasis can be induced centrally via Selective Serotonin-Reuptake Inhibitor (SSRI) antidepressants and Lysergic Acid Diethylamide (LSD), as well as N-Methyl-D-Aspartate (NMDA) glutamate antagonists, such as ketamine, in addition to opioid withdrawal and alcohol poisoning (Doughty, 2001; Foster, 1891).

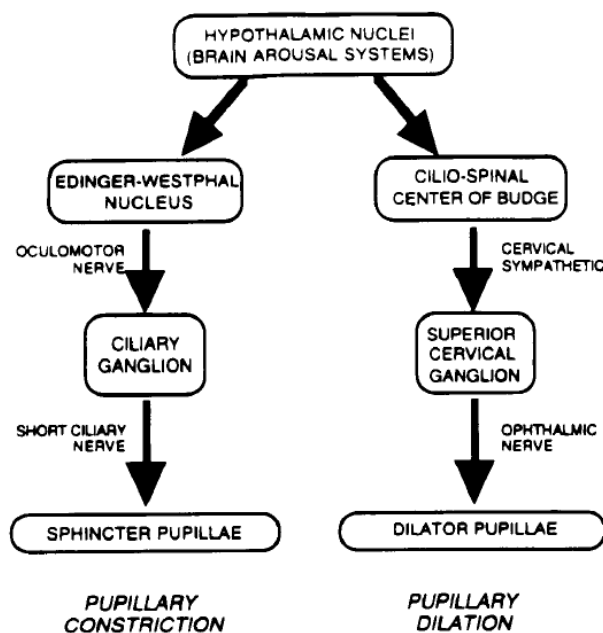


Figure 1-4: Pupillary constriction and dilation pathways (from Beatty, & Lucero-Wagoner, 2000).

As there are a number of stages in the efferent pathways, and cholinergic transmission is common to both sympathetic and parasympathetic pathways within the central nervous system, some substances – such as nicotine (Loewenfeld, 1999), antidepressants (Doughty, 2001) and alcohol (Skoglund, 1943) – can cause either constriction or dilation depending on concentration and activity in other parts of the pathway.

1.1.1.2. Pathology

Certain health states and pathologies can affect the size of the pupil. Locally pupil dilation can result from the increased intraocular pressure in glaucoma, and constriction can occur in response to treatment when the pressure decreases (Miller, & Newman, 2005). Neurological conditions affecting the central and peripheral nervous system, such as epilepsy, migraine, multiple sclerosis and lesions in the autonomic nervous system, can all have transient or long-lasting effects on both pupil-size and pupil responses (Grunberger, Linzmayer, Majda, Reitner, & Walter, 1996; Harle, Wolffsohn, & Evans, 2005). Sympathetic lesions, such as those that occur in Horner's Syndrome, can result in pupil constriction, while diabetes mellitus can lead to both sympathetic and parasympathetic autonomic neuropathy, resulting in loss of the light reflex (Argyll-Robertson pupil), and smaller resting pupil-size (Reeves, & Swenson, 2004). Argyll-Robertson pupil can also be caused by tertiary-stage neurosyphilis (Reeves, & Swenson, 2004). Compared to healthy populations, pupil abnormalities, such as increased or decreased diameter, decreased reactivity and pupil asymmetry, are also found in patients with organic and functional mental health problems such as Alzheimer's, depression and schizophrenia (Granholm et al., 2003; Sokolski, Nguyen, & DeMet, 2000; Steinhauer, Hakerem, & Spring, 1979), neurodiverse conditions like autism and ADHD (Martineau et al., 2011; Zahn, Little, &

Wender, 1978), and substance issues, including alcohol and heroin misuse (Ghodse, Greaves, & Lynch, 1999; Grunberger et al., 1998).

1.1.2. Pupil-Size Change

Löwenstein and Loewenfeld (1962) argue convincingly that all physical stimuli arriving at the senses, all somatic and visceral afferents, all mental processes including intentional efforts and motor responses, emotions, and all centrally mediated arousal responses, trigger a pupillary reflex dilation, also called the psychosensory reflex (Foster, 1891; Hess, 1965; 1975; Loewenfeld, 1999; Löwenstein, & Loewenfeld, 1962). The exceptions to this are certain visual reflexes, such as changing focus from far to near objects, and increased light falling on the retina, which cause constriction (see section 1.1.2.1). Some relatively early experiments in the late nineteenth century demonstrated pupillary dilations, occurring without changes in blood pressure, in response to peripheral tactile and pain stimuli, in animals that were conscious, under anaesthesia, and partially or completely paralysed with the acetylcholine antagonist curare (Schiff, & Foa, 1874; Schiff, 1875). Löwenstein and Friedman (1942) report that Schiff's earlier work (~1867) referred to the pupil as the body's "finest esthesiometer" (device measuring the skin's tactile sensitivity) (p. 969).

Interestingly, internal psychological events also cause pupil dilation and the first cognitive pupillometry study in humans was probably conducted by Heinrich (1896) who measured pupillary dilations evoked by mental multiplication. Such findings lead Oswald Bumke to observe in 1911 that: "...every active intellectual process, every psychical effort, every exertion of attention, every active mental image, regardless of content, particularly every affect just as truly produces pupil enlargement..." (cited by Hess, 1975, pp. 23-4). Löwenstein and Loewenfeld (1962) suggest that unlike the pupil-size changes that occur in response to increased peripheral activity in the

autonomic nervous system, the psychosensory reflex is under higher cortical control and reflects the level of cortico-thalamo-hypothalamic activity, which itself is influenced by sensory stimulation, emotions and spontaneous thought (Kahneman, 1973). There are extensive cortical and limbic inputs to both the Edinger-Westphal nucleus, increasing inhibition of the sphincter muscle, and to the hypothalamus, causing the dilator muscle to contract (Silk et al., 2009). In addition, the anterior cingulate cortex, thought to be involved in emotion regulation (Szabadi, & Bradshaw, 1996), inputs directly to the midbrain reticular formation which, when stimulated, increases pupil-size (Beatty, 1986). Pupil dilation also results from direct stimulation of limbic structures such as the amygdala (Koikegami, & Yoshida, 1953). However, despite higher order influences on pupil-size, even the earliest writings suggested that the pupil itself is not under voluntary control: "...it is not in our power to bring the will to act directly on the iris by itself. This fact alone indicates that the nervous mechanism of the pupil is of a special character..." (Foster, 1891, p. 1172).

Pupil-size is under the antagonistic control of both parasympathetic and sympathetic inputs, and each pathway is subject to various types of inhibition and excitation at each synapse. Therefore numerous internal and external factors, including stimulus characteristics like illumination, colour, contrast, and duration, are known to have individual and interacting effects on pupil-size (Yamaji, Hirata, & Usui, 2000). The following section briefly describes the major sources of pupil-size change, starting with changes in illumination via the light reflex, and accommodation via the near reflex. It is important to note that even these relatively rapid and automatic reflex processes are also modified by individual factors such as age, fatigue and emotional state, and these interactions will also be discussed.

1.1.2.1. *The Light Reflex*

The pupil is usually between 2-6mm in ambient illumination, dilating in dim light, and constricting in bright light (Beatty, & Lucero-Wagoner, 2000; Reeves, & Swenson, 2004; Young, & Biersdorf, 1954). This constriction is known as the Pupillary Light Reflex (PLR), and serves to regulate the total intensity of light entering the eye, optimising image quality (Kardon, 1995). In response to large increases in luminance the pupil can decrease in diameter by more than 50% in just 200ms (Miller, & Newman, 2005), peaking between 500-1000ms (Beatty, & Lucero-Wagoner, 2000). PLR amplitude and constriction velocity decrease with age, due to sympathetic deficit and smaller initial pupil-size (Birren et al., 1950; Bitsios et al., 1996a).

Photosensitive retinal ganglion cells respond to increased light falling on the retina. The PLR acts afferently via the optic nerve to the midbrain pretectal region, and efferently to the Edinger-Westphal nucleus, then via parasympathetic fibres in the left and right oculomotor nerves to the ciliary ganglions causing contraction of the pupil sphincter muscle. Simultaneous inhibition of the sympathetic fibres innervating the dilator muscles causes the antagonistic muscles to relax, and the pupil constricts to reduce the amount of light falling on the retina (Loewenfeld, 1999). Both pupils generally change equally even if light only enters one eye, unlike some animals, such as frogs and birds, where pupil light reflexes are independent (Foster, 1891). In humans this is due to the equal bilateral input from each retina to the pretectal region, and bilateral output from the pretectal region to the Edinger-Westphal nuclei (Reeves, & Swenson, 2004; Thompson, 1947).

Although the PLR is automatic, its amplitude can be reduced by evoking a simultaneous psychosensory dilation with emotional or painful stimuli (Bender, 1933; Gang, 1945; Miller, & Newman, 2005). For example, Bitsios, Szabadi and Bradshaw

(1996b; 2002) found that when a light stimulus follows the threat of an aversive stimulus, such as electric shock, the reflexive constriction is smaller (fear-inhibited light reflex), and is negatively correlated with self-reported anxiety. This amplitude reduction can itself be attenuated by anxiolytic drugs such as diazepam (Bitsios, Philpott, Langley, Bradshaw, & Szabadi, 1999). Sustained cognitive processing during a task (such as counting backwards in intervals of 7) has also been shown to diminish the PLR, an effect that may reflect cortical inhibition of the Edinger-Westphal nuclei (Steinhauer, Condray, & Kasperek, 2000; Steinhauer, Siegle, Condray, & Pless, 2004). In both the fear- and cognitive task-inhibited PLR overall pupil-size is larger to begin with, due to increased emotional arousal and cognitive load (see section 1.2). Consequently, careful consideration needs to be given to how pupil-size change is measured and whether absolute changes in PLR amplitude should be of the same magnitude when initial diameters vary (see Chapter 2, section 2.1.2.2, for further discussion of this issue in pupillometry research).

1.1.2.2. The Darkness Reflex

The darkness reflex involves dilation of the pupil due to cessation of the sympathetic inhibition caused by a constant light source (Löwenstein, & Loewenfeld, 1964).

Pupils take longer (300ms) to begin dilating in response to darkness than to constrict to light, and this latency does not change with age. However the maximum velocity of constriction and maximum dilation, of between 3-9mm, both decrease from about age fifteen (Birren et al., 1950; Bitsios et al., 1996a; Miller, & Newman, 2005). The amplitude of the darkness reflex is related to the length of time in the dark (Stark, 1962, cited by Loewenfeld, 1999), and rather than just an absence of the PLR, it is thought that signals from the retina may inhibit the oculomotor nerves. Retinal disease, which leads to the loss of these signals, causes “paradoxical” constriction in response to darkness (Miller, & Newman, 2005).

1.1.2.3. *The Lid-Closure Reflex*

The lid-closure reflex manifests as a systematic miosis (of around 8% surface area), occurring after blinks, which is followed by redilation (DeLaunay, 1949; Hupé, Lamirel, & Lorenceau, 2009; Loewenfeld, 1999), although some researchers have observed only a dilation after eye-blinks (Fukuda, Stern, Brown, & Russo, 2005).

Photosensitive retinal cells increase in sensitivity during brief (<500ms) interruptions of light, such as during a blink, however the lid-closure reflex does not occur after blinks in darkness (Hupé et al., 2009).

1.1.2.4. *The Accommodation Response*

The pupil constricts when we change our focus from looking at a far away object to a near object, and when the eyes converge, such as when looking at the tip of the nose (Foster, 1891; Löwenstein, & Loewenfeld, 1964; Reeves, & Swenson, 2004). The purpose of the accommodation response is to maintain a focussed image on the surface of the retina. Unlike a camera, where the lens moves forward to focus on a nearer object, the lens of the eye is elastic; it changes in thickness, becoming more convex, and therefore increases in refractive power (Beatty, & Lucero-Wagoner, 2000; Foster, 1891; see Figure 1-5). Young healthy eyes are able to accommodate within 350ms (Erichsen, Hodos, & Evinger, 2000), adjusting focus from distant to near objects via three muscle groups – lens curvature is increased by contraction of the ciliary muscle/release of the zonule fibres, the eyes converge through contraction of the medial rectus muscle, and the pupillary sphincter muscle constricts the pupil (Beatty, & Lucero-Wagoner, 2000).

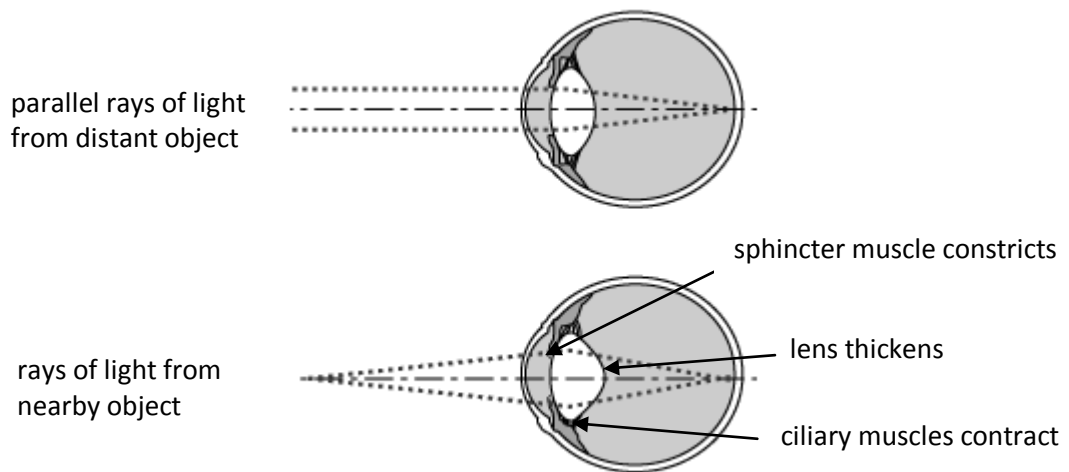


Figure 1-5: The accommodation reflex – the pupil constricts as the lens thickens (adapted from Nave, 2010).

The pupil constriction in the near reflex shares the same efferent parasympathetic pathway with the PLR from the Edinger-Westphal nucleus onwards, and so is also susceptible to central inhibition from concurrent psychosensory dilations (Miller, & Newman, 2005). Its magnitude is also influenced by illumination, such that in dim light the constriction associated with focussing on a near object is smaller than in bright conditions (Miller, & Newman, 2005). Accommodation, convergence and PLR constrictions are neurologically distinct and are dissociated in conditions like Argyll-Robertson pupil (where there is no PLR but normal responses to accommodation and convergence), diphtheritic neuritis (where there is a preserved PLR but no change with accommodation), and pretectal lesions (where normal constriction occurs for accommodation, but not for convergence; Reeves, & Swenson, 2004; Spector, 1990).

1.1.2.5. Pupillary Hippus

The iris is a vascular structure and rhythmic changes in pupil-size of around 1% occur with heart beat and breathing due to fluctuations in blood pressure (Foster, 1891). However, there are other rhythmic but irregular, oscillating, consensual (therefore of central origin) contractions and dilations of reasonably large amplitude ~1mm (10-

20%) with a period of between 5-25s (0.04-0.2Hz; Bouma, & Baghuis, 1971; McLaren, Erie, & Brubaker, 1992; Woodmansee, 1966). They occur under constant illumination and fixation, and vary with intensity of illumination rather than pulse and respiration (Loewenfeld, 1958). This pupillary unrest is also known as ‘hippus’ (possibly from *hippos*, Greek for ‘horse’, suggestive of a galloping rhythm; Beatty, & Lucero-Wagoner, 2000), which was originally used to describe pathological changes in pupil-size that can occur in phase with EEG recordings in seizure disorders (Müller-Jensen, & Hagenah, 1978) or respiration in Cheyne-Stokes (Sullivan, Manfredi, & Behnke, 1968). However, spontaneous consensual hippus usually has no clinical significance, it is induced by changes in lighting level, becoming more obvious in brighter light and when the pupil is small (Bouma, & Baghuis, 1971; Miller, & Newman, 2005).

In addition to sensory and endogenous influences, hippus varies according to “arousal” and “cognitive effort” (see sections 1.2.2 and 1.2.2.1). Hippus is amplified by fatigue and passivity, but is suppressed by alertness and carrying out mental activities such as mental arithmetic (Bouma, & Baghuis, 1971; Kahneman, 1971; Miller, & Newman, 2005). Due to this suppression, a problem arises in taking baseline measures of pupil-size before pupillometry experiments in that hippus may be occurring during the baseline period, whereas it will be attenuated during the experimental task, resulting in a hippus artefact in pupil-size data (Janisse, 1977; see Chapter 2 section 2.1.2.2 for further discussion).

1.1.2.6. *Iris Colour*

Iris colour has been considered a possible confound in pupillometry research. Dark irises are associated with decreased sympathetic reactivity, muscle motility, and contraction amplitude compared to paler irises (Beck, 1967; Dain, Cassimaty, &

Psarakis, 2004; Gambill, Ogle, & Kearns, 1967; Hess, 1975; Spector, 1990).

However, not all research into iris colour has found an effect on pupil-size (Birren et al., 1950; Bradley et al., 2010; Goodrich, 1974; Kumnick, 1954; Wenger, & Videbeck, 1969) and Janisse (1977) suggests that previous findings may have been an artefact of distinguishing pupil-size from a dark iris. The mixed findings may reflect the numerous morphological and chemical factors which influence perceived iris colour (with >240 degrees of freedom; Daugman, 2003) and, for these reasons, this variable is not given further consideration.

1.1.2.7. Role of the Locus Coeruleus in Stimulus-Evoked Dilations

The Locus Coeruleus (LC) is a neuromodulatory nucleus in the dorsal pons of the brainstem which responds to salient stimuli (e.g., targets) with a transient increase in firing rate. The LC projects throughout the forebrain, providing all of the forebrain and most of the brain's Norepinephrine (NE). This influence makes the LC partly responsible for regulating all cognitive, emotional and motivational states (see Berridge, & Waterhouse, 2003, for a review; Samuels, & Szabadi, 2008a; Sara, 2009). The LC receives afferents from the Anterior Cingulate Cortex (ACC) and Orbitofrontal Cortex (OFC), which Aston-Jones and Cohen (2005) have argued are involved in monitoring "task-related utility" (the cost/benefit of continuing with the current task vs. looking for a new opportunity), and supplying information about conflict and reward in the cognitive system (Aston-Jones et al., 2002; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Rajkowski, Lu, Zhu, Cohen, & Aston-Jones, 2000; Zhu, Iba, Rajkowski, & Aston-Jones, 2004). The LC-NE system also plays a role in wakefulness (Aston-Jones, Foote, & Bloom, 1984; Jouvet, 1969), simple decision-making and the regulation of task engagement through distribution of attentional resources (e.g., Aston-Jones, Rajkowski, & Kubiak, 1997; Clayton, Rajkowski, Cohen, & Aston-Jones, 2004), and recent research has argued that it also

plays a major role in the modulation of the psychosensory pupillary response (Gilzenrat, 2006).

NE is the neurotransmitter released by the sympathetic fibres that innervate the pupil dilator muscle. Single-cell intracranial recordings and in-vivo stimulation of the LC in conscious monkeys and rodents performing memory and sensory tasks show two modes of activity – regular, continuous tonic firing (1-3Hz), interspersed with short bursts of phasic firing (8-10Hz; Aston-Jones, Chiang, & Alexinsky, 1991; Aston-Jones et al., 1997; Aston-Jones, Rajkowski, Kubiak, & Alexinsky, 1994). Both modes have been shown to vary with vigilance and task performance measures, such as stimulus processing efficiency (Aston-Jones, & Cohen, 2005; Rajkowski, Majczynski, Clayton, & Aston-Jones, 2004; Samuels, & Szabadi, 2008b; Usher, Cohen, Servan-Schreiber, Rajkowski, & Aston-Jones, 1999), and a high correlation (0.6) between spike frequency and pupil diameter has been found, whereby large pupil diameter equates to high LC activity (Rajkowski, Kubiak, & Aston-Jones, 1993; 1994; see Figure 1-6).

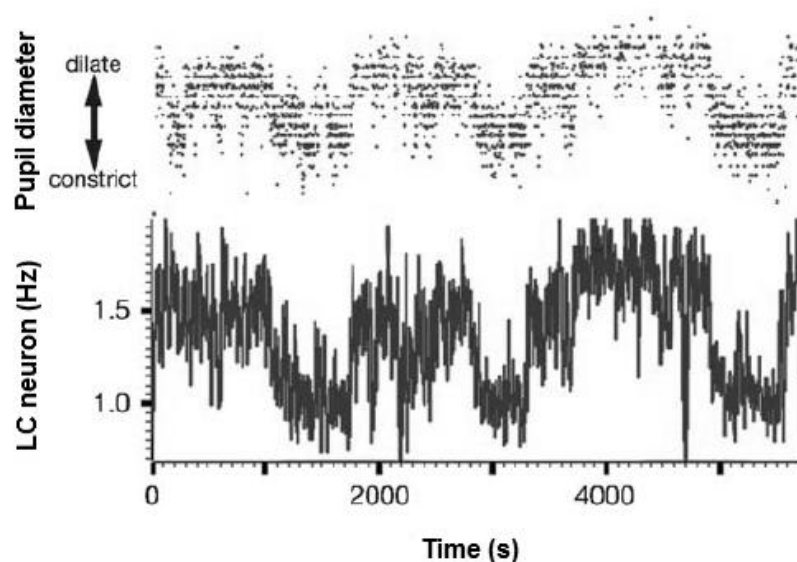


Figure 1-6: Concurrent pupil-size and LC neuron recording in the monkey (from Rajkowski, Kubiak, & Aston-Jones, 1993).

A differentiation between tonic and phasic pupil activity has been observed in humans (Dureman, & Scholander, 1962), and whilst no current technique allows direct recording of LC neurons in humans, the relationship between pupil-size and task performance found in other animals has been confirmed in human participants (Gilzenrat, Cohen, Rajkowski, & Aston-Jones, 2003). Additionally the stimulus-evoked pupil dilations seen in human vigilance experiments (e.g., Beatty, 1982a) are consistent with the phasic pupil dilations which arise from phasic LC activity reported in the animal literature (Aston-Jones, & Cohen, 2005; Murphy, Robertson, Balsters, & O'Connell, 2011). Even though the relationship between pupil-size and LC activity has not been fully characterised, and is subject to debate (see Nieuwenhuis, De Geus, & Aston-Jones, 2011a), pharmacological up- or down-regulation of central NE release in humans provides strong supporting evidence by mimicking increased and decreased LC activity respectively. For example, sympathomimetic drugs such as modafinil, yohimbine, and reboxetine increase subjective alertness by increasing central NE; they also increase baseline pupil-size, reduce pupillary variability and fatigue waves, and reduce the amplitude and velocity of the darkness reflex (Hou, Freeman, Langley, Szabadi, & Bradshaw, 2005; Phillips, Bitsios, Szabadi, & Bradshaw, 2000a; Phillips, Szabadi, & Bradshaw, 2000b). In contrast, sympatholytic drugs such as clonidine and prazosin decrease subjective alertness by decreasing central NE; they also decrease baseline pupil-size, increase pupillary variability and fatigue waves, and increase the velocity of the darkness reflex (Hou et al., 2005; Phillips et al., 2000a; 2000b). As a result of this converging evidence, researchers increasingly use pupil-size as an indirect indicator of LC activity to investigate aspects of human attention, such as orienting of attention to external cues (Gabay, Pertzov, & Henik, 2011), switching attentional focus (Einhauser, Stout, Koch, & Carter, 2008), and changes in cognitive control state (Gilzenrat, Nieuwenhuis, Jepma, & Cohen, 2010; Jepma, & Nieuwenhuis, 2011).

Connection between the LC and the pupil can also be seen in patient populations. The LC receives input from the vagus nerve, known to play a role in memory formation (Clark, Krahl, Smith, & Jensen, 1995; Clark, Naritoku, Smith, Browning, & Jensen, 1999), via projections from the solitary tract. Individuals who benefit from vagus nerve stimulation (which modulates NE release), such as those with treatment-resistant epilepsy, depression, anxiety disorders, Alzheimer's Disease (AD), migraine, fibromyalgia, and tinnitus (for a review see Groves, & Brown, 2005; Engineer et al., 2011; Ghanem, & Early, 2006; Lange et al., 2011), have been shown to demonstrate atypical pupil responses. For example, migraine sufferers show various pupil abnormalities such as anisocoria (Evans, & Jacobson, 2003), reduced velocity and amplitude of the PLR (Mylus, Braune, & Schepelmann, 2003), and hyper- or hypo-responsivity to pharmacological agents that affect the autonomic system (Fanciullacci, Galli, Pietrini, & Sicuteri, 1977). Harle et al. (2005) showed that anisocoria and inter-ocular differences in PLR latency persist during the non-headache phase, independent of time since last migraine, severity or frequency, suggesting sustained autonomic imbalance in migraineurs. In AD patients, topical application of tropicamide, a cholinergic antagonist, produces significantly larger pupil dilations than for vascular dementia patients or young non-AD patients (Iijima et al., 2003), and smaller peak PLR constriction amplitude (Granholm et al., 2003).

1.2. Pupillometry Research Literature

Pupillometry has been used to look at a wide variety of psychosensory and physiological functions in a variety of animals (cats, chickens, dogs, fish, frogs, guinea pigs, monkeys, pigeons, rabbits, and rats), producing a literature that has grown exponentially. In a fascinating dissertation, Loewenfeld (1958) reviewed over 1300 pupil references dating back to the 1st century AD. She made a distinction between 114 references to the pupil in historical pre-1830 literature, and 1204 pieces

from 1830-1957, because the existence of muscles within the iris, and their role in the dilation of the pupil, was only established around this time. By the time of her later publication, Loewenfeld (1999) listed over 15,000 references to pupil research leading up to ~1985 (Steinhauer, 2002).

Psychosensory fluctuations in pupil-size are usually no more than 0.5-1.0mm and are therefore difficult to see with the naked eye (Beatty, 1982b; Beatty, & Lucero-Wagoner, 2000; methods of measuring pupil-size in cognitive pupillometry are described in Chapter 2). These tiny yet consistent pupillary changes have no apparent functional purpose or evolutionary cost, and appear to reflect dynamic changes in cognitive processing (Beatty, & Lucero-Wagoner, 2000). Although the pupil is under peripheral autonomic control, without an obvious link to central processing, evidence shows that the variations reliably and precisely track changes in cognition (Beatty, 1982b; 1986; Goldwater, 1972). Described as “a permanently implanted electrode” and “the only visible part of the brain” (Janisse, 1977, p. 1), the pupil is of particular interest to cognitive psychophysiology researchers because it potentially provides a unique physiological reporter variable to measure psychological processes independently of subjective report (Beatty, & Lucero-Wagoner, 2000).

Task-evoked pupil-size changes begin 400ms post-stimulus (Partala, & Surakka, 2003), peaking after around 1-2s, and constricting once the task is complete, either slowly (Kahneman, & Beatty, 1966; Hess, 1972) or rapidly (Bernhardt, Dabbs, & Riad, 1996) dependant on post-processing. Like other psychophysiological measures, such as blood flow in functional Magnetic Resonance Imaging (fMRI), it has been argued that pupillary changes provide an *indirect* measure of processing intensity without causal links or face validity (Just, & Carpenter, 1993). As such they are also subject to the problem of ‘psychophysiological inference’ (the assumption that a

physiological response has a consistent one-to-one and context-independent relationship with the psychological variable of interest) originally raised by William James (1890) (Cacioppo, & Tassinari, 1990; Cacioppo, Tassinari, & Berntson, 2000). However, Beatty and Lucero-Wagoner (2000) liken the use of Task-Evoked Pupillary Responses (TEPRs) to the use of reporter genes in molecular biology, which have advanced understanding of the genome, and suggest that TEPRs might do the same in psychophysiology for understanding cognition.

Given the vast number of factors influencing pupil-size, and the numerous afferent pathways involved, the question is whether it is possible to isolate and study the systematic effects of individual influences within the pupillary signal. The next section of this chapter will briefly review the main areas of human cognitive pupillometry from the 1960s onwards.

1.2.1. Eckhard H. Hess

Pupil changes in relation to cognitive activity were first demonstrated around the turn of the twentieth century (e.g., Heinrich, 1896; Mentz, 1895; Roubinovitch, 1900), but the findings remained largely within the European literature. It was not until the 1960s, and the experiments (and controversy) of Eckhard H. Hess, that a resurgence of interest in pupil-size in North America led to more systematic and thorough investigation, using increasingly sophisticated pupillometry techniques and equipment (see Beatty, 1982b; Hess, 1975; Kahneman, 1973; Steinhauer, 2002). Although by no means the first, Hess is widely acknowledged as a key figure in the history of cognitive pupillometry (called pupillography prior to the use of computerised measures), establishing a clear relationship between psychological processes and changes in pupil-size during three decades of published research (Janisse, 1977).

In their first study, Hess and Polt (1960) showed four male and two female participants, pictures of a baby, mother and child, partially nude male, partially nude female, and a landscape. Pupil-size was manually measured from a 16mm filmstrip, of the participants' eyes during the task, projected onto a screen. Male participants' pupils dilated most to the picture of the nude female, whereas female participants' pupils dilated most to the picture of mother and child. Hess and Polt (1960) replicated their results, interpreting the findings as showing the interest value of the pictures, whereby more interesting stimuli elicited larger pupil dilations. In 1964 Hess and Polt conducted the first rigorous investigation of pupil dilation in relation to mental arithmetic, confirming the findings of Heinrich (1896). Again, measuring pupil-size from projected 16mm filmstrip, they asked five participants to carry out four mental multiplications, giving a total of twenty data points. They found that pupil-size increased whilst the answer was being calculated, and maximum dilation increased from 10.8% to 21.6% approximately monotonically in proportion to calculation difficulty (see Figure 1-7), suggesting that changes in pupil-size could be used to directly measure cognitive activity as it occurs (Hess, & Polt, 1964).

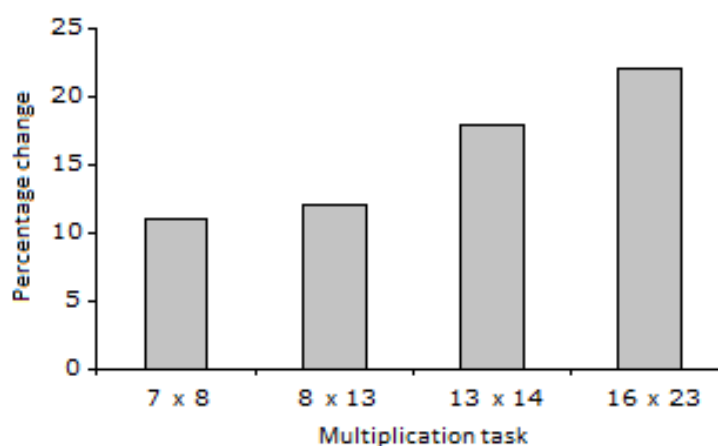


Figure 1-7: Percentage pupil-size change in relation to baseline pupil diameter (in mm, not reported), shows an almost perfect monotonic increase of pupil-size with task difficulty (adapted from Hess, & Polt, 1964).

The finding that pupil dilation increases with increasing multiplication difficulty has since been replicated by other researchers (Bradshaw, 1968a; Klingner, 2010; Marshall, 2002; Payne, Perry, & Harasymiw, 1968). Ahern and Beatty (1979, 1981) showed that the TEPR was smaller for more intelligent college students than for less intelligent counterparts when carrying out the same arithmetic problems. The same was true for digit span and sentence comprehension, suggesting more efficient information processing in the participants of higher psychometric intelligence, who required fewer cognitive resources (Ahern, & Beatty, 1979; 1981). However, Beatty and Lucero-Wagoner (2000) observe that this research did not address the possible role of practice and over-learning in the higher intelligence group.

Much of Hess' later research was around the concept of "emotional valence" and his controversial aversion-constriction hypothesis: the pupil dilates to positive-affect stimuli and constricts to negative-affect stimuli (Hess, 1965; Hess, & Polt, 1960). Hess (1972) showed participants affectively loaded photographs of crippled children or mutilation, and reported an initial dilation followed by constriction caused by the "shock value" of the stimuli. Other researchers have used the idea of emotional valence to infer people's attitudes and preferences from their pupil-size. For example, it was shown that participants' pupils would dilate to images of preferred political leaders and candidates, and constrict to undesirable images (Barlow, 1969; 1970; Clark, & Ertas 1975; Hess, 1965).

1.2.1.1. Criticisms of Hess and Early Work

Hess has been criticised by independent researchers and reviewers (e.g., Dooley, & Lehr, 1967; Goldwater, 1972; Hakerem, 1973; Janisse, 1973; Mueller, 1970; Peavler, & McLaughlin, 1967; Woodmansee, 1966; Zuckerman, 1971) for repeatedly using very small sample sizes, not reporting all potentially relevant results, not using

evidence to back up his interpretations or conclusions, citing “unpublished” pilot data, pooling his data across experiments, rarely presenting appropriate statistical analyses (or none at all before 1966), imprecise methods, providing insufficient detail for replication, and claiming to have discovered things that other people had already published first and with better methodology (Janisse, 1977; Löwenstein, & Loewenfeld, 1962). The fact that Hess’ (1975) literature review ignores any research disagreeing with his own findings has made it difficult to draw any conclusions from Hess’ bulky and ambiguous literature (Janisse, 1977). Describing Hess’ work as “inane twaddle”, eminent and world-renowned pupil researcher Irene Loewenfeld (1999) dismissed most of the emotion-based pupil research, not in objection to the existence of the well-established psychosensory reflex, but because she considered the research methodology and analyses to be flawed, requiring replication with more appropriate and rigorous techniques. Loewenfeld also passionately denounced Hess’ aversion-constriction theory due to her intimate knowledge of pupil physiology.

Hess answered his critics by stating that aversion-constriction was present only for “certain” people and stimuli, rather than all people and all aversive stimuli. The aversion-constriction hypothesis has not been replicated by other researchers using carefully controlled studies (e.g., White, & Maltzman, 1978; Pavio, & Simpson, 1966; Schaefer, Ferguson, Klein, & Rawson, 1968). It has instead been criticised on methodological grounds in several literature reviews, which suggest instead that constriction is the result of habituation or decreased interest in the experiment (arousal decrement) (e.g., Woodmansee, 1966), and that the brief stimulus-evoked phasic dilations are more likely to be the result of cognitive activation, whereas emotional arousal effects are longer lasting and more likely to influence tonic or baseline pupil-size (e.g., Beatty, 1982; Beatty, & Lucero-Wagoner, 2000; Goldwater, 1972; Janisse, 1977). There is plenty of evidence supporting psychosensory dilation,

but no good evidence for aversion-constriction (Janisse, 1977). Pupillometry has particular difficulties, some of which are unique among psychophysiological techniques (Janisse, 1977). Critics suggest that Hess did not control the luminance of his stimuli or his laboratory, and that “aversive” stimuli were brighter, causing constriction of the pupil through the PLR rather than an emotional response (Janisse, 1977). Beatty (1972) suggests precautions researchers that can take to equate stimulus brightness and contrast; however this is much harder when conducting pupillometry studies in naturalistic situations (Wang, 2010).

Due to the PLR, it is imperative that researchers control both background illumination, for example constant artificial lighting, and the global and local luminance of the stimuli, especially when presented on a computer screen (Janisse, 1977). This second point has been the source of considerable controversy (e.g., Hess, Beaver, & Shrouf, 1975; Janisse, 1973; Loewenfeld, 1966; Woodmansee, 1966). Pictorial and photographic stimuli have been strongly criticised because they vary greatly in luminance both globally between stimuli and locally within different regions of a single stimulus, which can create artifactual pupil-size changes (e.g., Goldwater, 1972; Janisse, 1977; Woodmansee, 1970; Zuckerman, 1971). The light and dark properties of images can generate pupil-size changes of around the same magnitude as psychosensory changes (Janisse, 1977) and should be taken into consideration when employing visual scanning (Pomplun, & Sunkara, 2003; Van Orden, Limbert, Makeig, & Jung, 2001), visual search (Backs, & Walrath, 1992; Porter, Trościanko, & Gilchrist, 2007) or photographs (Dabbs, & Milun, 1999; Libby, Lacey, & Lacey, 1973). It is not only Hess who has been criticised for failing to control luminance – many of the early studies using pictorial stimuli did not take this into consideration, inaccurately reporting pupil constriction in response to the affective quality of the stimuli (e.g., Tanck, & Robbins, 1970). Peavler and McLaughlin (1967) showed four female, and

four male participants, pictures including three of clothed females and one of a nude female. They found that participants' pupils only dilated to the nude picture, and constricted to the images of clothed females; however the clothed pictures were brighter than the others.

Tryon (1975) surveyed twenty possible sources of variation and confounds in pupillometry, many of which have been mentioned above. When participants are presented with multiple stimuli in an experiment, adaptation or habituation can occur, where the overall diameter of the pupil decreases, the magnitude of the TEPRs decrease, and the speed of contraction increases (Löwenstein, & Loewenfeld, 1952; Lehr, & Bergum, 1966; Tryon, 1975). Goodrich (1974; 1975) highlights the fact that actual pupil-size is distorted by the cornea, which has a lens power equivalent to 38-48 dioptres (Janisse, 1977). Conducting experiments using participants under the age of 30 (Woodmansee, 1966), having stimuli 3-4m from participants (Hakerem, & Sutton, 1964), and using relatively short trials, all help to reduce pupil-size variation due to the near-vision reflex, which occurs when participants lose or change focus due to age, fatigue or boredom (Janisse, 1977). Even when luminance is held constant, other visual features such as spatial frequency, patterns and movement can also influence pupil-size (Barbur, Wolf, & Lennie, 1998; Nakayama, Yasuike, & Shimizu, 1990; Sahraie, & Barbur, 1997; Sooter, & van Norren, 1980; Ukai, 1985). Due to a phenomenon known as the pupillomotor Purkinje effect, the pupil dilates more in response to coloured stimuli (chromatic) than grey-scale stimuli (achromatic), and constricts more to shorter wavelengths as luminance increases (Bouma, 1962; Kohn, & Clynes, 1969). In addition, colours can have emotional meaning (Bouma, 1962; Kohn, & Clynes, 1969; Miller, 1967). For these reasons, it is highly likely that a visual stimulus will produce a change in pupil-size in a cognitive pupillometry study.

Important and more rigorous programs investigating psychological influences on pupil-size have been carried out by researchers such as Kahneman, and Paivio and Simpson (Janisse, 1977).

1.2.2. Arousal

The term “arousal” is vague, somewhat contentious, and is used by different researchers to refer to a variety of constructs, such as emotions, sexual attraction, or attention, each of which has a large and overlapping literature (see Staal, 2004; Neiss, 1988). The most popularized pupillometry arousal research has involved sex, racism or fear; for example, using pictures of nudes (Lawless, & Wake, 1968), people of different races (Woodmansee, 1967), or fear of electric shock (Polt, 1970). Stimuli were designed to elevate participants’ mental and physical arousal, stimulating the sympathetic nervous system and the release of adrenaline into the blood stream, leading to pupil dilation.

Many experiments (e.g., Aboyoun, & Dabbs, 1998; Bull, & Shead, 1979; Hess, 1965; Hicks, Reaney, & Hill, 1967; White, & Maltzman, 1978) have shown that pupil-size is linearly related to the level of sexual arousal (Janisse, 1977). Janisse (1977) asks whether pupillary dilation accompanying “Don Juan[’s] ... statements of undying devotion” (p. 11) is due to sexual arousal (Zuckerman, 1971) or to the fact that he is lying (Bradley, & Janisse, 1975, cited in Janisse, 1975). Using erotic and suspense films, Bernick, Kling and Borowitz (1971) were able to show that pupil-size may discriminate sexual arousal from more generalised arousal. However, as Janisse (1977) has indicated, most of the pupillometric research around sexual arousal used pictorial or video stimuli, varying in luminance (e.g., Hess, Seltzer, & Shlien, 1965; Nunnally, Knott, Duchnowski, & Parker, 1967; Peavler, & McLaughlin, 1967; Scott, Wells, Wood, & Morgan, 1967; Lawless, & Wake, 1969). An additional potential

confound is that at that time in North America nude pictures were more novel than clothed pictures, and the pupil also responds to novelty (Andreassi, 2000).

Since Hess' original studies, and contrary to the aversion-constriction hypothesis, other researchers have since shown that the pupil dilates to emotional stimuli of both positive and negative valence (Janisse, & Peavler, 1974; Stelmack, & Mandelzys, 1975). For example, Guinan (1967) showed that average pupil-size was larger for high emotionality words than low emotionality words, particularly in the first 2.5s, and suggested that the emotional content of the stimuli was causing autonomic arousal. The same has been found for acoustic stimuli by Partala and Surakka (2003) who played participants ten positive, ten negative and ten neutral sounds, finding that their pupils dilated more to the positive and negative sounds compared to the neutral sounds (0.2mm vs. 0.14mm). However other researchers have found a U-shaped function of pupil-size where the pupil is larger for neutral stimuli than for slightly positive and negative stimuli (e.g., Gunther, & Lussier, 1975, cited in Janisse, 1977).

Urry et al. (2006) presented negative and neutral affective images and asked participants to intentionally increase their emotional response to the stimuli (imagining the situation happening to them), decrease their response (viewing the situation as fake), or simply attend to the stimuli in a control condition. They showed that actively enhancing emotional responses increased initial and sustained pupil dilation compared to decreasing emotional responses, and that in both emotional regulation conditions pupil-size was larger than in the unregulated 'control' condition. However there were multiple influences on pupil-size including cognitive effort, imagery and emotional arousal, whereas the 'attend' condition did not involve a cognitive/imagination task and so was insufficient to act as control. More concerning

is the fact that the stimuli were colour photographs, therefore potentially introducing visual confounds.

Further evidence for the link between emotion and pupil-size comes from studies showing that depressed participants (medicated and unmedicated) show different pupil responses to emotional stimuli compared to non-depressed participants. For example, Siegle, Steinhauer, Carter, Ramel and Thase (2003a) found increased and sustained pupil dilation in depressives compared to controls for up to 30s when identifying the valence of emotional words. The extent of pupil-size increase also correlated with self-reported rumination, suggesting that it reflected sustained elaboration of emotional information processing in depressive participants (Siegle et al., 2003a; Siegle, Granholm, Ingram, & Matt, 2001; Siegle, Steinhauer, Carter, & Thase, submitted; Siegle, Steinhauer, & Thase, 2004).

A substantial subset of the arousal literature has focussed on the relationship between pupil-size and state and trait anxiety, which lead to larger pupil-sizes consistent with sympathetic arousal; a larger resting pupil diameter in perpetually anxious participants was noted by Bumke as early as 1903 (cited in Janisse, 1977). The influence of anxiety on pupil-size is so consistent and dramatic that Paivio and Simpson (1966), and Kahneman (1973) discussed the potential confounding factor of anxiety in pupillometry experiments designed to measure cognitive effort or task difficulty. Carver (1971) and Johnson (1971) proposed that progressive pupil dilation with increasing task difficulty occurred as a result of increased anxiety and emotional arousal due to anticipation, because in many studies (e.g., Kahneman, & Beatty, 1966) participants were informed in advance about the increasing difficulty. However Peavler (1974) found difficulty-related increases in the absence of prior knowledge of the task, while Simpson and Molloy (1971) showed that task-related and difficulty-

related increases in pupil-size persist even in anxious participants, who show larger baseline pupil-sizes and response-related dilations than non-anxious participants when expected to make an overt response. This is due to factors such as performance anxiety, anticipation of making a response, explicitly making a decision, anxiety about making a mistake, being judged, or receiving feedback based on a response (Paivio, & Simpson, 1966; Simpson, 1969; Simpson, & Molloy, 1971; Simpson, & Paivio, 1968). However anxious participants had qualitatively the same cognitive load pupil-response curve as in other experiments manipulating task difficulty (Kahneman, 1973; Paivio, 1973). Even in non-anxious participants, the offering of incentives and penalties (and therefore the introduction of an element of risk) increases TEPRs (Kahneman, Peavler, & Onuska, 1968b; Kahneman, & Peavler, 1969).

Painful or 'startling' stimuli, such as heat or loud noise, cause dilation with a latency of 300-500ms (Janisse, 1977; Loewenfeld, 1958; Nunnally et al., 1967). Chapman, Oka, Bradshaw, Jacobson and Donaldson (1999) provided "noxious" electrical fingertip stimulation at four intensities, increasing from faint to almost unbearable and found that peak pupil-size rose with intensity. Polt (1970) showed that even the threat of an electric shock caused pupil dilation, but this could also have been the result of increased cognitive effort to answer correctly under the threat of shock. However, most of the arousal research conducted in the 1960s and 1970s was subject to the same criticisms as Hess' emotion research. There was no established common way of reporting or analysing results, luminance confounds were common, and little communication or shared learning of methodological issues occurred between researchers (Janisse, 1977).

1.2.2.1. *Fatigue & Sleepiness*

Pupil-size also reflects decreases in arousal, and pupil changes are associated with the sleep-wake cycle (Löwenstein, & Loewenfeld, 1964). Diameter is largest when individuals are well-rested, decreasing with fatigue and reaching its smallest diameter prior to sleep (Foster, 1891; Loewenfeld, 1999). Fatigue also increases amplitude and frequency of pupillary hippus, particularly in darkness (Löwenstein, & Loewenfeld, 1952; 1964; Yoss, Moyer, & Hollenhorst, 1970). Pathological sleep deprivation, such as in Excessive Daytime Sleepiness and Narcolepsy, causes distinctive patterns of pupillary movements for narcoleptics compared to healthy participants, and for treated versus untreated narcoleptics, both in light and darkness. For example, narcoleptics have smaller baseline pupil-sizes and show less random pupillary noise (as opposed to spontaneous changes caused by hippus) compared to healthy participants (O'Neill, Oroujeh, Keegan, & Merritt, 1996; O'Neill, Oroujeh, & Merritt, 1998; Pressman et al., 1984; Yoss, 1970; Yoss, Moyer, & Ogle, 1969). Narcoleptic participants also demonstrate differences on cognitive pupil measures, for example working memory overload occurs after storing fewer digits; however earlier and faster pupil dilation compensates for smaller baseline pupil-size to arrive at a peak dilation equivalent to controls (O'Neill, & Trick, 2001). Parkinson's Disease is often accompanied by arousal symptoms such as sleepiness, which are positively correlated with pupillary unrest (Jain et al., 2011).

Using techniques developed by Yoss (1969; 1970), variation in fatigue-related pupil changes have the potential to be useful in monitoring alertness in professions where vigilance is critical, such as drivers (Recarte, & Nunes, 2003; Recarte et al., 2008; Walzl, Hagen, & Prummer, 2007; Yoss, 1969), pilots (Dehais, Causse, & Pastor, 2008; Yoss, Moyer, Carter, & Evans, 1970), naval vessel operators (de Greef, Lafeber, van Oostendorp, & Lindenberg, 2009), industrial and construction workers

(Geacintov, & Peavler, 1974; Wilhelm et al., 2010) doctors (Wilhelm, Widmann, Durst, Heine, & Otto, 2009) and telephone operators (Geacintov, & Peavler, 1974). It also points to a promising avenue of pupil research in monitoring attentiveness and/or affect in real world scenarios through Human-Computer Interaction (HCI) technology (e.g., Bailey, & Iqbal, 2008; Iqbal et al., 2004; Lin, Imamiya, & Mao, 2008; Oliveira, Aula, & Russell, 2009; Rowe, Sibert, & Irwin, 1998). Pupil-size correlates with task difficulty during individual stages of a task, and decreases during transitions between tasks (Iqbal et al., 2004; Bailey, & Iqbal, 2008; Oliveira et al., 2009). Researchers have used changes in pupil-size to identify transitions between subsections of tasks when a user could be interrupted with the least amount of disruption (Bailey, Busbey, & Iqbal, 2007; Iqbal, Adamczyk, Zheng, & Bailey, 2005). Another aim of HCI is to enable computers to establish a user's affective state (e.g., Barreto, Zhai, Rishé, & Gao, 2007; Lanatà, Armato, Valenza, & Scilingo, 2011). Gao, Barreto and Adjouadi (2010) developed an algorithm which was able to identify stressed states in participants with 77.8% accuracy using pupil-size, Galvanic Skin Response (GSR) and pulse measurements.

1.2.3. “Cognitive Effort” – Kahneman

The majority of pupil research from the 1960s onwards concerns the relationship between pupil-size and “cognitive effort”, a complex concept which some have described as the proportion of available attentional resources assigned to a task (see Cain, 2007; Moray, 1979). Other related concepts include “mental workload” and “processing load” due to the fact that cognitive resources are limited (Miller, 1956), and performing two tasks simultaneously usually leads to decreased performance on one or both (Beatty, 1982b; Kahneman, 1973). A key player in cognitive effort research was Nobel Prize winner Daniel Kahneman, whom Janisse (1977) credits

with the methodical exploration of the concepts of cognitive “loading” and “unloading”, “processing”, “mental effort” and “rehearsal” within pupillometric research.

Kahneman (1973) argued that general “arousal” can be seen as a response to task demands, and therefore “mental effort” is part of arousal. However, he emphasised the necessity for researchers to distinguish between dilations caused by processing load and dilations caused by other elements of arousal, such as muscle activity or anxiety, in rigorously designed experiments with careful consideration of potential confounds (Kahneman, & Wright, 1971). Pupil dilations are largest during task performance, compared to before and after, and those corresponding to correct responses are usually larger than those for failures (Kahneman, 1973). In addition, studies have shown that although behavioural responses have a small effect on pupil-size, due to performance anxiety and muscular exertion (see section 1.2.5), dilation occurs due to task performance even in the absence of an overt response, so these other factors cannot account for the majority of the pupil-size change (Kahneman et al., 1968b).

Kahneman (1973) describes a useful physiological index of mental effort as one which is responsive to variation within-task, faithfully tracking changes in participant effort whilst they carry out the task, between-tasks, identifying which tasks are more difficult and therefore require more effort, and between-participants, showing that different people invest different amounts of effort in a task. Beatty (1982b) says that pupil responses meet all three of these criteria (Beatty, 1982b). An interesting example of between-participant differences is that of “intelligence”; however pupillometric studies show mixed findings, with some researchers finding no differences in pupil-size for participants of high and low intelligence (e.g., Bernick, Altman, & Mintz, 1972; Daly, 1966; Simpson, & Molloy, 1971), whilst others have

found larger pupil-sizes for participants with lower IQ (e.g., Ahern, & Beatty, 1979; 1981; Verney, Granholm, Marshall, Malcarne, & Saccuzzo, 2005), or, in contrast, larger pupil-sizes for participants with higher IQ (e.g., Boersma, Wilton, Barham, & Muir, 1970). Verney et al. (2005), for example, controlled stimulus brightness and found that more intelligent participants performed better at a visual backwards masking task, which evoked smaller pupil responses, compared to less intelligent participants who performed worse (detected fewer targets, allocated more attentional resources to non-target stimuli) and had larger pupil-sizes. If more intelligent participants perform better on tasks because they find tasks easier than less intelligent participants, it might be expected that their pupils would be smaller; however if they perform better because they exert more effort, it might be expected that their pupils would be larger than less intelligent participants (Janisse, 1977). The question becomes whether pupil-size measures difficulty or effort, and whether effort and perception of difficulty vary with intelligence (Janisse, 1977). As dilation varies with effort, which varies with intelligence and cognitive resources, then individuals will vary on the amount of effort required to carry out the same task *and* the amount of effort-related dilation, but the two may be indistinguishable.

Kuc and Janisse (1967; 1976, cited by Janisse, 1977) compared successful and unsuccessful digit span recall at a 50% difficulty threshold (50% of trials were correct at that level of difficulty), and measured intelligence using a specific measure (Digit Span Forward subscale of WAIS, 1955) following completion of the task. Using this approach they found that pupil-size was larger for correct trials and suggested this was due to increased mental effort leading to success, since task difficulty was held constant, and subjective difficulty was associated with incorrect trials which had smaller pupils. They also found no significant main effect of intelligence overall, although there was a trend towards larger pupil-size during loading for the high

intelligence group, who also gave more correct answers (Janisse, 1977). Potential confounds included the larger number of verbal responses associated with correct trials than incorrect trials, and confidence, although Janisse (1977) concludes that pupil-size changes are far more likely to reflect effort than difficulty and that using “intelligence” as a factor in pupillometry studies is not straightforward. However, using measures that are task-specific, rather than general, and making comparisons of correct and incorrect trials may simplify interpretation of results and lead to comparisons of more appropriate groups (Janisse, 1977). Daly (1966, cited by Janisse, 1977) suggested that fluctuations in pupil-size, which decrease when participants concentrate, might be a better measure of problem-solving efficiency than maximum pupil-size (Kahneman, 1971; Janisse, 1977).

Other aspects of task performance such as accuracy, motivation and memory load have all been linked to pupil-size changes (Janisse, 1977). Pupil-size has repeatedly been shown to increase with increasing mental effort in a variety of tasks (see Table 1; for reviews see Beatty, 1982b; Beatty, & Lucero-Wagoner, 2000; Goldwater, 1972; Janisse, 1977). Kahneman (1973) asks whether pupil-size changes that occur in response to different types of task can be legitimately compared in terms of the amount of effort expended. For example, measurement at a single time point may not represent total effort for tasks involving sustained effort (Kahneman, 1973). Tasks involving rapidly decaying short term memory, such as digit-span or pitch-discrimination, or tasks requiring participants to respond quickly to stimuli, generate both time-pressure and large pupil-sizes (Kahneman, 1973). ‘Difficulty’ manipulations do not increase task difficulty equally between different types of requirement, for example there is a larger gap between easy and difficult arithmetical problems, due to storage and rehearsal, than between easy and difficult sentence comprehension problems (Elshtain, & Schaefer, 1968; Kahneman, 1973).

Task	Key Pupil-Size Findings	Selected Authors
Signal detection	Pupil response only present when participants reported flash, not when identical flash went unreported; large response to low probability or omitted stimuli, similar to P300	Light: Hakerem, & Sutton, 1966; tones: Beatty, 1975
Pitch discrimination	Max increased monotonically as comparison & reference tones became closer in pitch, and therefore harder to discriminate	Kahneman, & Beatty, 1967; Schlemmer, Kulke, Kuchinke, & Van Der Meer, 2005
Sentence comprehension	Complex sentences induced larger pupillary responses than simpler sentences (0.25mm vs. 0.21mm) and increased latency to peak by 116ms	Just, & Carpenter, 1993; Wright, & Kahneman, 1971
Digit recall	Increases incrementally during loading, returning to baseline during recall; peak increased with string length; overload occurs sooner in schizophrenic participants	Beatty, 1966; Beatty, & Kahneman, 1966; Granholm et al., 1997; Kahneman, & Beatty, 1966; Simpson, & Hale, 1969
Dual task performance	Larger for dual-task than single visual search, but similar for single digit transformation; during dual-task, errors increase & decrease with pupil-size, suggesting maximal processing capacity reached	Kahneman, Beatty, & Pollack, 1967
Incentive	Larger for high than low incentives	Kahneman, & Peavler, 1969; Kahneman, Peavler, & Onuska, 1968
Random motor responses	larger when participant chose to move than when instructed	Simpson, & Hale, 1969
Visual search	Larger for more difficult searches; increases as task progresses (memory load?)	Pomplun, & Sunkara, 2003; Porter, Trościński, & Gilchrist, 2007
Spatial ability	Increased with angular disparity when judging irregular hexagons; greater dilations for low than high spatial ability participants	Just, & Carpenter, 1995
Processing speed	Increased at 75% & 100% max processing speed capacity, constriction at 125%	Poock, 1973
Deception	Larger for guilty/lying than innocent/truthful participants	Berrien, & Huntington, 1943; Dionisio, Granholm, Hillix, & Perrine, 2001

Table 1: Selection of tasks used to investigate the relationship between pupil-size and “mental effort”.

Some researchers have found larger pupil-sizes during tasks that intuitively feel ‘easy’ than for tasks we might consider more difficult. For example, dilations to paired-associate recall are 4-6 times larger than during learning (Kahneman, & Peavler,

1969); and the pupil dilates more whilst retaining five digits for immediate recall, which most participants perform with 'ease', than whilst attempting to listen to and comprehend a long complex message (Carver, 1971). Large dilations also accompany the prompted recall of over-learned personal information such as age or phone number, which should be 'easy' to retrieve (Beatty, & Kahneman, 1966; Kahneman, 1973; Schaefer et al., 1968). It may be misleading to conclude that more effort is required for recall than for learning, or that recall is more 'difficult', when the task demands are different (Kahneman, 1973).

Steinhauer et al. (2004) used a novel approach to look at the sympathetic and parasympathetic contributions to pupil-size during sustained cognitive effort. Participants performed an easy task (add 1) and difficult task (subtract 7) in normal light and complete darkness. Their pupils dilated equally to both tasks in darkness, but dilated more to the hard task in the light. Steinhauer et al. (2004) suggested that in addition to sympathetic dilation, there was also parasympathetic inhibition of constriction for the difficult task in the light, allowing larger dilations to occur. They repeated the experiment having used eye-drops to selectively block the parasympathetic sphincter muscle (tropicamide), or the sympathetic dilator muscles (dapiprazole), or neither (placebo). Dilation was seen in each condition, however effects of task demands and light condition on pupil-size, and the previously seen interaction between the two, were only present when dapiprazole was used to block the sympathetic dilator muscles. It was absent when tropicamide was used to block the parasympathetic constrictions, suggesting that it is increased parasympathetic inhibition of pupil constriction that leads to larger pupil dilation during more difficult tasks, whereas sympathetic activity is less differentially affected (Steinhauer et al., 2004). Direct cortical input, indirect cortico-thalamic-hypothalamic input and arousal-related reticular pathway activity have all been linked to inhibition of the Edinger-

Westphal region through which the pupillary sphincter muscle is controlled (Bonvallet, & Zbrozyna, 1963; Löwenstein, 1955).

1.2.3.1. Signal Detection & Discrimination

One of the most straightforward cognitive tasks is asking a participant to indicate when they detect a near-threshold sensory stimulus, for example, a weak flash of light in a darkened room. Signal detection as applied to pupillometry was first reported by Hakerem and Sutton (1966), who recorded tiny pupil responses ($<0.1\text{mm}$) to brief flashes of light near the visual threshold followed by a tone. Participants were dark adapted for 20 minutes prior to the experiment and asked to press a button after the tone if they had seen the light (Hakerem, & Sutton, 1966). Button presses were counterbalanced across experiments to control for response-related dilations, and participants experienced pupil dilations to light stimuli that were too weak to evoke a light-reflex. Interestingly the stimulus-evoked response was only present in trials where participants reported seeing the flash, not in trials where an identical flash went unreported. The relationship between pupil-size and conscious awareness is returned to in the general discussion of Chapter 3, section 3.4.

In a standard auditory signal detection experiment Beatty and Wagoner (1975; 1976) asked participants to detect 100ms 1kHz tones in a white noise background and press one of four buttons (yes-certain, yes-uncertain, no-uncertain, and no-certain). In a 2 (signal: present, absent) x 2 (decision: yes, no) x 2 (confidence: certain, uncertain) design, they found that pupil-size was the same for the conditions where the signal was absent. For the signal present conditions, pupil changes were related to stimulus-response category, whereby yes-certain decisions evoked the largest dilations, followed by yes-uncertain decisions, no-uncertain decisions, and no-certain decisions. Beatty and Wagoner concluded that pupil-size changes were only related

to outcome in the presence of a signal. The findings support a cognitive or effort-based, rather than anxiety or emotion-based, interpretation of pupillary dilation since the two 'uncertain' types of trial evoked an intermediate sized response (Beatty, & Lucero-Wagoner, 2000; Janisse, 1977).

Discriminating between two signals, presented either simultaneously or in succession, is more complex and resource intensive than signal detection, leading to larger overall pupil dilation in signal discrimination than in signal detection (Beatty, & Lucero-Wagoner, 2000). For example, maximum pupil-size increases monotonically as a comparison and reference tone become closer in pitch, making them harder to discriminate (Kahneman, & Beatty, 1967).

1.2.3.2. *Working Memory*

In a series of paced auditory serial recall/digit span tasks Kahneman and Beatty (1966) presented participants with digit strings of 3-7 items at a rate of one per second (loading phase), asking them to recall the items in the same order at the same rate (unloading phase). During loading pupil-size increased incrementally with each successive item, due to increasing rehearsal in working memory (Baddeley, & Hitch, 1974; Kahneman, & Beatty, 1966), and peak size increased with increasing string length (0.1mm vs. 0.55mm for 3 and 7 digits). This was followed in the unloading phase by a brief dilation (approximately one second *before* unloading began), then decrements as items were recalled, due to cessation of rehearsal (Johnson, 1971), until pupil-size returned to baseline at the end of the trial.

Kahneman and Beatty (1966) replicated the loading and unloading effect by asking participants to add 1 to each digit (e.g., hear 3-9-1-6, report 4-0-2-7), and in a different task, with 4-item strings of words (high-frequency monosyllabic nouns), and found that the increased task difficulty lead to larger pupil dilations, particularly for the

digit transformation. Kahneman and Beatty's (1966) classic experiments have been replicated using modern remote eye-tracking techniques and the same results have been found (Klingner, 2010).

Peavler (1974) investigated the effects of saturating central processing resources by using strings of up to 13 digits in a paced recall task and found that "overloading" did not lead to further increases in pupil-size. After increases in response to the participants' maximal information processing capacity (Miller, 1956; Poock, 1973) of 7-8 digits, the pupil-size remained just below the maximum until recall (Peaver, 1974). Other investigators have found that pupil-size starts to decline once demands outstrip available resources, particularly in populations with cognitive impairments such as in schizophrenia (Granholm, Asarnow, Sarkin, & Dykes, 1996; Granholm, Morris, Sarkin, Asarnow, & Jeste, 1997), suggesting that differences arose from variation in instructions – Peavler (1974) asked participants to try their best, so they may have actively maintained their maximum digit strings through rehearsal (Granholm et al., 1996; 1997). These studies also contribute compelling evidence that it is cognitive, rather than emotional or anxiety-related changes, which cause the pupil-size increases as failure-related emotional responses should cause dilation to peak following overload (Beatty, & Lucero-Wagoner, 2000).

1.2.3.3. Visual Search

Increasing pupil-size has been demonstrated in visual search, which involves working memory (e.g., Porter et al., 2007). By comparing counting and searching within the same arrays, Porter et al. (2007) found that counting involved sustained pupil dilation throughout the task, whereas search involved little dilation to begin with and increasing dilation throughout the task as spatial memory demands increased. Manipulating search difficulty by increasing the number and variety of distracters, they

found that more difficult searches, and larger or more complex spatial arrays, produced larger pupil-sizes than easier searches and smaller or less complex arrays.

1.2.3.4. Effort and the Red Pupillary Reflex

The fundus or “red” reflex occurs when light shone into the eye is reflected by the ocular fundus, and is seen most commonly in everyday life as “red eye” in photographs. Kruger (1975; 1977a; 1977b) found a 10% increase in red reflex luminance accompanying increases in arithmetic task difficulty, which he attributed to an increase in accommodation (Kruger, 1977a). He confirmed this by using cycloplegic drugs (which paralyse the ciliary muscles and prevent accommodation), and found that luminance was no longer related to task difficulty (Kruger, 1977b). However, cognitive effort-related pupil-size increase is a potential confound as this would cause a brighter red reflex by allowing more light into the eye. Cycloplegic drugs cause mydriasis as well as accommodative paralysis, which would also explain the loss of relationship between red reflex and task difficulty; other researchers have found that accommodative response remains constant or even decreases as cognitive processing increases (Jainta, Hoormann, & Jaschinski, 2008; Davies, Wolffsohn, & Gilmartin, 2005).

1.2.3.5. Language & Comprehension

There is a robust literature linking pupil-size to variations in the complexity of language processing, such as speech perception (Zekveld, Kramer, & Festen, 2011), language translation (Hyönä, Tommola, & Alaja, 1995), letter perception (Beatty, & Wagoner, 1978), syntax (Schluroff, 1982) and semantic processing (Hyönä et al., 1995). For example, increased grammatical complexity and ambiguity increases processing load, short-term memory demands and decision time, producing larger peak dilations and longer peak latencies than simpler sentences (Just, & Carpenter,

1993; Schluroff, 1982; Schluroff et al., 1986; Stanners, Headley, & Clark, 1972; Wright, & Kahneman, 1971). Beatty and Wagoner (1978) asked participants to decide whether letter pairs were physically the same (AA, aa), phonetically the same (Aa), or from the same category (vowels: Ae). Larger pupil-sizes were found to phonetic pairs than physical pairs, and larger pupil-sizes to vowel pairs than other pairs, concluding that pupil-size reflected the amount of processing needed to decide whether letters were the same (Beatty, & Lucero-Wagoner, 2000). However, researchers have found that changes in cognitive load associated with reading texts of varying difficulty is not reflected in pupil-size (Schultheis, & Jameson, 2004; Iqbal, Zheng, & Bailey, 2004), perhaps due to minor differences between easy and difficulty reading tasks.

1.2.3.6. *Imagery*

Paivio and Simpson (1966) asked participants to create images in their mind of concrete and abstract nouns. Measuring manually from projected 16mm film strip, they found that when an overt response was required, abstract words evoked significantly larger pupil-sizes with longer latency and duration compared to concrete words, and that both abstract and concrete words evoked larger pupil-sizes than control slides of equivalent brightness. This finding has been replicated several times (Colman, & Paivio, 1969; Paivio, & Simpson, 1968; Simpson, & Climan, 1971; Simpson, Molloy, Hale, & Climan, 1968; Simpson, & Paivio, 1966; 1968) and interpreted as abstract words being more difficult to imagine, and therefore taking longer and requiring more cognitive effort (Kahneman, 1973).

1.2.3.7. *Lie Detection*

Despite being under autonomic control, relatively little research has been conducted into the activity of the pupil during lying (Janisse, 1977). The majority of studies have

concluded that deception-related changes in pupil-size are related to additional cognitive effort involved in fabricating answers (e.g., Dionisio, Granholm, Hillix, & Perrine, 2001) and/or increased anxiety related to concealing information and the fear of being caught (e.g., Berrien, & Huntington, 1943; Bradley, & Janisse, 1975, cited in Janisse, 1977). Several studies have involved participants committing a mock crime and hiding it from the experimenter (e.g., Lubow, & Fein, 1996), and only a couple have investigated pupil-size whilst asking participants to lie in a memory paradigm (Dionisio et al., 2001; Heaver, & Hutton, 2011). Detailed discussion of pupillometry studies involving lie-detection can be found in Chapter 5.

1.2.3.8. Auditory Stimuli

Importantly, cognitive influences on pupil-size are not dependant on the visual modality, a number of studies have shown that these effects also occur for auditory stimuli. Klingner (2010) replicated three classic cognitive pupillometry studies (digit span, mental arithmetic, and vigilance) using the same stimuli presented auditorally and visually under conditions of equivalent brightness and contrast. In all three, pupil-size showed qualitatively the same pattern of dilations with increasing task difficulty in both modalities, however pupil-size was on average almost twice as large for auditory stimuli as for visual stimuli (Klingner, 2010). The qualitative similarity in results suggests that pupil responses reflect post-perception task demands, and that there is increased processing load for auditory tasks, whereas visual presentation facilitates comprehension and processing (Klingner, 2010).

1.2.4. Attention

“Attention” remains one of the most enigmatic concepts in cognitive psychophysiology (Beatty, & Lucero-Wagoner, 2000) and definitions made over 100 years ago remain

relevant today: “the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought.

Focalisation, concentration, of consciousness are of its essence” (James, 1890, pp. 403-4). The influence of attention on pupil-size is “widely accepted” (Janisse, 1977, p. 2), and changes in pupil-size appear to closely chart central attentional resources (Beatty, & Lucero-Wagoner, 2000). Researchers have related pupillary dilation observations with attention-related concepts such as “interest” and “novelty” as part of the autonomic orienting response to salient stimuli (Lynn, 1966; Nieuwenhuis et al., 2011a; Pavlov, 1927; Sokolov, 1963). Libby et al. (1973) noted that interesting or attention-getting pictures evoke the largest pupil-sizes, and Pratt (1970) reported that more complex and interesting (less predictable) shapes produce larger pupil dilations. Over repeated exposure, the effect of novel, unfamiliar or unpredictable stimuli lessens (Pratt, 1970).

However, as discussed in section 1.2.1.1, pictorial stimuli confound pupil-size studies. An alternative method of assessing attention effects without visual confounds is to use auditory stimuli. Beatty (1988) recorded pupil-size whilst participants were instructed to detect target tones presented to one ear and to ignore all tones presented to the other ear. Similar to findings in the ERP literature (e.g., Hink, Van Voorhis, & Hillyard, 1977), he found no detectable increase in pupil-size to tones in the unattended channel, tiny (0.015mm) dilations to non-targets in the attended channel and large stimulus-evoked dilations to targets (Beatty, 1988). In an auditory vigilance task Beatty (1982a) asked participants to sustain attention to a task and detect infrequent target tones amongst non-target tones for 48 minutes. As the task progressed, behavioural and pupil-size data demonstrated a vigilance decrement, as participants became less sensitive to targets, more conservative in their judgement criteria, and experienced smaller stimulus-evoked dilations.

When focused attention increases, sympathetic activity increases (sympathetic dominance) and parasympathetic activity decreases, however, there is often only a small correlation between markers of sympathetic dominance, such as pupil diameter, GSR, blood pressure and heart rate (Kahneman, 1973). Autonomic markers often respond differently to stimuli, for example the largest pupil-sizes and slowest heart rates may occur to the most interesting pictures in a selection, referred to by Lacey (1967; Libby et al., 1973) as “directional fractionation”.

1.2.4.1. Locus Coeruleus

As discussed in section 1.1.2.7, the activity of the Locus Coeruleus (LC) has been demonstrated to change with pupil-size, level of arousal and attention to a task. The LC is thought to be involved in the flexible allocation of cognitive control as a function of internal and external requirements (Cohen, Botvinick, & Carter, 2000; Cohen, Aston-Jones, & Gilzenrat, 2004). The Adaptive Gain Theory (AGT) of LC-NE function (Aston-Jones, & Cohen, 2005) states that phasic LC activity optimises task engagement for exploitation of a known source of reward, and that tonic activity promotes disengagement to allow exploration for new sources. This exploit-explore balance is important in the cognitive control and adaptive regulation of behaviour, and the model has enabled the successful prediction of pupil response in tasks involving target detection, conflict and reward (Gilzenrat, 2006). Very recently baseline pupil-size has been shown to correspond to performance dynamics and task engagement in humans, whereby increases in prestimulus pupil-size are associated with decreasing task engagement, and decreases in prestimulus pupil-size correspond linearly with increases in task-evoked dilations (phasic responses) and task engagement (Gilzenrat et al., 2010; Jepma, & Nieuwenhuis, 2011).

Gilzenrat et al. (2010) used an auditory target detection task to show that on a trial-by-trial basis baseline pupil-size accurately predicted engagement and disengagement from task performance according to AGT. Manipulation of conflict and reward within tasks caused participants to either persist with the task or choose to give up, and this behaviour was preceded by increases and decreases in baseline pupil-size respectively. Smaller baseline pupils were inversely followed by larger stimulus-evoked dilations, as described by the phasic mode of the LC. Gilzenrat et al. (2010) provided the first evidence of this sort, and felt that pupil-size could be used to index control state, even though in this experiment LC activity was inferred from pupil dynamics rather than direct recording of neuronal responses.

Nieuwenhuis, Gilzenrat, Holmes and Cohen (2005b) extended the LC-NE model to describe the “attentional blink” (Raymond, Shapiro, & Arnell, 1992), a temporary inability to perceive the second target when two targets are presented temporally very close together. The authors proposed that this impairment results from the neural refractory period of the LC-NE system as observed in monkeys after a phasic response to a target (e.g., Usher et al., 1999). The P300 ERP component is also suppressed for the second target under the same experimental conditions, particularly when the P300 to the first target is large, leading Nieuwenhuis et al. (2005b) to suggest that it also reflects LC-NE phasic activity,

1.2.4.2. Blinks

As discussed in section 1.1.2.3, after a blink the pupil briefly contracts before redilating (DeLaunay, 1949). A number of studies have now shown a link between cognition and spontaneous blinks, which has lead to some researchers excluding post-blink periods from their analyses (e.g., Hupé et al., 2009). Studies have shown longer blink suppression and higher blink rates accompanying high cognitive load

than low cognitive load, and shorter blink latencies for positive compared to negative stimuli (e.g., Ohira, 1996; Ohira, Winton, & Oyama, 1998; Recarte, Perez, Conchillo, & Nunes, 2008). Siegle, Ichikawa and Steinhauer (2008) suggest that because they occur under cortical control, blinks distinguish separate stages of information processing, relating to changes in cognitive states or resource allocation, and regulating the time in which information is acquired (Fogarty, & Stern, 1989; Stern, Walrath, & Goldstein, 1984). Siegle et al. (2008) proposed that blinks complement pupil-size changes in the analysis of cognitive tasks.

1.2.4.3. Decision Making & Uncertainty

The act of making a simple decision in a laboratory setting leads to a significant increase in pupil-size compared to being instructed (Simpson, & Hale, 1969). Most pupillometric studies require participants to respond, therefore unless the variable under investigation is 'choice', task instructions need to be very clear so as to remove decisions and uncertainty from requirements (Janisse, 1977). As stimulus uncertainty increases, baseline pupil-size increases and phasic pupil amplitude decreases (Bradshaw, 1968b; 1969; Levine, 1969, cited in Hakerem, 1974; Pratt, 1970). With the exception of Simpson (1969), the literature shows a sympathetic-like dilation in anticipation of parts of each trial, such as onsets, offsets, changing stimuli, or response prompts (Bradshaw, 1969; Janisse, 1977; Nunnally et al., 1967).

1.2.5. Physical Effort

The pupil dilates in response to physical effort (Foster, 1891; Hakerem, & Sutton, 1966; Hupé et al., 2009; Nunnally et al., 1967). Nunnally et al. (1967) asked participants to lift weights of increasing then decreasing mass for ten seconds each, with a ten second rest in between, whilst measuring pupil-size. They found that mean

pupil-size was larger during the lifting period than the following rest period, and that as the mass of the weight increased and decreased so did pupil-size. Nunnally et al. (1967) reported pilot results showing the same relationship between pupil-size and fist clenching exercises, and concluded that that pupil-size is related to muscle tension (see Figure 1-8).

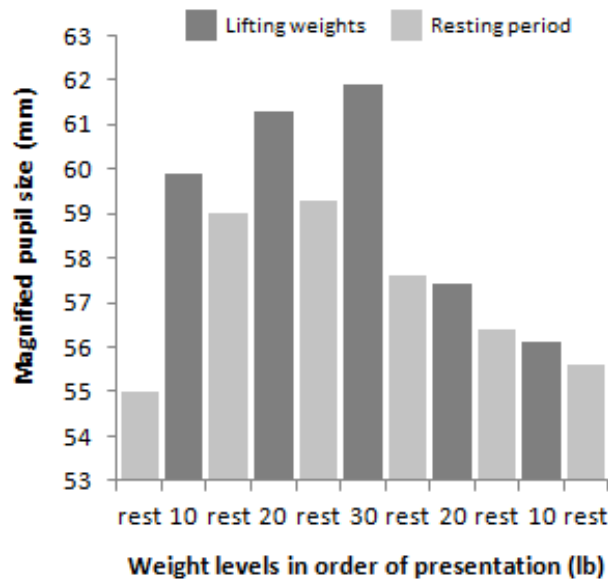


Figure 1-8: Mean pupil-size (magnified x17.5) in relation to 10 second lifting periods and intervening 10 second resting periods (adapted from Nunnally, Knott, Duchnowski, & Parker, 1967).

Participants do not need to lift large weights for motor activity to influence their pupil-size, and, in addition to task demands and performance anxiety, an overt response may induce pupil-size changes through muscle tension from moving the mouth or a finger (Paivio, & Simpson, 1966; Richer, & Beatty, 1985; Simpson, 1969; Simpson, & Molloy, 1971; Simpson, & Paivio, 1968). Hupé et al. (2009) showed participants ambiguous moving stimuli and asked them to press a button when their bistable percept changed. They found that 70% of the 5% average change in pupil-size area was due to making the motor response.

Despite the fact that overt responses cannot account for all pupil-size changes, some researchers have shown that pupil responses only occur when participants are required to make an overt response to the task (Hakerem, & Sutton, 1966; Paivio, & Simpson, 1966; Simpson, & Paivio, 1968). For example, in Hakerem and Sutton's (1966) signal detection task, when participants were not required to indicate when they had detected a flash, pupil-size did not increase to any of the flashes, whereas when a decision was required, pupil-size increased when flashes were detected. Counterbalancing the conditions in which a response was required ensured that the difference was not due purely to motor effort, and Hakerem and Sutton suggested the results reflected a higher level of vigilance in the reporting conditions. It is not difficult to imagine that participants may be less engaged with the task if their performance is not being externally assessed. However, Simpson and Paivio (1966) used an imagery task to show that larger pupil dilations occurred to abstract words than concrete words even without a motor response. As this is in contrast with findings by the same authors that the difference only occurred when a response was required (Paivio, & Simpson, 1966; Simpson, & Paivio, 1968), it highlights the importance and likely contribution of task instructions and continued participant motivation.

1.2.6. Pupil-Size and Concurrent Psychophysiological Measures

Pupil measures may be especially suited to experimental studies because like other psychophysiological markers, pupil-size can covertly and continuously monitor the time course of cognitive processes. Its millisecond time resolution enables researchers to observe the dynamics of processing demands with minimal latency, and data collection is not dependent on participant response (Hyönä et al., 1995; Kramer, 1991). To give a fuller picture, changes in pupil-size can be combined with

other psychophysiological markers of cognitive effort such as heart rate variance (Lin et al., 2008; Rowe et al., 1998), functional Magnetic Resonance Imaging (fMRI; Just, Carpenter, & Miyake, 2003), Electroencephalography (EEG; Schacter, 1977), Event-Related Potentials (ERPs; Just et al., 2003; Kok, 1997), plasma 17-hydroxycorticosteroid levels (Bernick et al., 1971), Positron Emission Tomography (PET; Just, Carpenter, Keller, Eddy, & Thulborn, 1996), GSR (Colman, & Paivio, 1969; Kahneman et al., 1969), and other eye-tracking measures such as blink rate (Recarte et al., 2008; Siegle et al., 2008), or fixation time (de Greef et al., 2009).

As it is under autonomic control, the pupil demonstrates characteristics such as arousal decrement (Woodmansee, 1966), parasympathetic rebound (Rubin, 1964, cited by Janisse, 1964), and habituation (Löwenstein, & Loewenfeld, 1962). Wilder's (1958) Law of Initial Values (LIV) states that for physiological responses, as the initial value gets higher so the response becomes smaller for enhancing stimuli and larger for reducing stimuli (Jin, 1992). However, despite sympathetic and parasympathetic input, at times the pupil does not respond in line with other peripheral autonomic efferents, such as heart rate. Examples of this "directional fractionation" (Lacey, 1959; 1967) include the effects of modafinil, which increases noradrenergic input to the pupil without influencing heart rate and salivation (Hou et al., 2005). Libby et al. (1973) found that changes in heart-rate and pupil-size share only 19% of their variance, and that there is great individual variation in correlation between the two variables (between -0.52 and 0.39, average 0.35). Pupil-size has a greater correlation with GSR of 0.30-0.50 (Colman, & Paivio, 1969; Scott et al., 1967), but Colman and Paivio (1969) suggest that pupil-size "may be a more sensitive peripheral response than GSR during cognitive tasks" (p. 296), for example, differentiating concrete from abstract imagery, and easy from difficult paired-associate learning (Colman, & Paivio, 1969; McElvain, 1970). It is difficult to establish from these

observations how much of this variation is due to pupil-size and GSR indexing related but different processes that sometimes co-occur, leading to a higher correlation under certain conditions, or whether they index the same processes but that pupil-size is just more sensitive.

In a paced serial recall/digit span task Kahneman et al. (1969) found that increases in GSR and pupil-size tracked increases in task difficulty, whereas heart rate decreased as difficulty increased. Of the three measures, pupil-size, measured with a “high-speed” (1Hz) infrared film camera, was the most consistent (Janisse, 1977). A negative correlation between pupil-size and heart-rate was also found by Kuc and Janisse (1967, cited by Janisse, 1977) in participants performing digit span under stress ($r = -.55$), replicating previous findings (Clark, 1975; Libby et al., 1973), whereas consistent with the views of Löwenstein and Loewenfeld (1962) there was a minimal positive relationship under low-stress conditions (0.14). This was because although pupil-size increased more during loading, overall pupil-size under both conditions was the same, yet heart-rate was faster throughout the high-stress condition than the low-stress condition. Additionally, the authors concluded that pupil-size reflected participant intelligence, correctness of answer and cognitive aspects of the task, whereas heart rate reflected emotional aspects (Janisse, 1977). The pupillary system is therefore a sensitive, specific and comparatively low-noise measure of psychophysiological changes (Beatty, & Lucero-Wagoner, 2000).

A growing number of studies have combined pupillometry with neuroimaging techniques such as ERP or fMRI (e.g., Brown, Kinderman, Siegle, Granholm, Wong, & Buxton, 1999; Conway, Jones, DeBruine, Little, & Sahraie, 2008; Friedman, Hakerem, Sutton, & Fleiss, 1973; Hawkes, & Stow, 1981; Just, & Carpenter, 1993; Just et al., 1996; 2003; Kuipers, & Thierry, 2011; Ledoux et al., 2010; Muller-Jensen,

& Hagenah, 1979; Murphy et al., 2011; Siegle, Steinhauer, Stenger, Konecky, & Carter, 2003b; Steinhauer, 1982; Van Droof et al., 2010). Techniques can be compared if researchers use the same paradigms (e.g., Vilberg, & Rugg, 2008), and pupil-size is easily acquired alongside other techniques, providing complementary data about the processes under investigation (Just et al., 2003).

Siegle et al. (2003b) recorded concurrent pupil-size and fMRI data whilst participants carried out a digit sorting task, and found activity in the middle frontal gyrus had a similar time-course to pupil-size. By recording pupil data on its own outside the fMRI scanner using the same participants, the researchers were able to show that the parametric increase with task difficulty was the same in both contexts. They were also able to use the individual variation in pupil-size to model the activity in the middle frontal gyrus to improve sensitivity and specificity, showing that pupil-size accurately reflects task-related cognitive activity as measured by fMRI (Siegle et al., 2003b). However, because pupil-size is measured continuously and tracks changes in task requirements with low latency (0.1-0.5s) it may be a more consistent measure of general cognitive effort than measures such as GSR (Kramer, 1991).

Just and colleagues (e.g., 1993; 1996; 2003) have investigated psychophysiological indices of working memory load for two decades, using a wide range of executive processing, language processing, spatial and memory task whilst recording ERPs, fMRI and pupillometry data. Just et al. (2003) reviewed the literature and showed that similarities exist between all three measures, for example, response magnitude during tasks. They concluded that when being used to measure the same paradigms the techniques tap the same common process, “capacity utilisation” or cognitive load.

1.2.7. Memory

Although working memory has received considerable attention (see section 1.2.3.1), very few researchers have studied the influence of Long-Term Memory (LTM) retrieval such as recognition on pupil-size. In the next section current models of recognition memory will be outlined, before returning to a discussion of pupil-size and recognition memory in section 1.3.2.3.

1.3. Recognition Memory

1.3.1. Models of Recognition Memory

Almost everything we do relies, more or less, on our capacity to learn, store and retrieve information from memory. However, the precise structure of human memory and the underlying cognitive processes are so complex, that even after centuries of research using increasingly sophisticated methods, the current literature still contains ongoing debate, and reports divided opinion on how we should model this fundamental function (e.g., Diana, Reder, Arndt, & Park, 2006; Dunn, 2004; Macmillan, & Rotello, 2006; Malmberg, Holden, & Shiffrin, 2004; Murdock, 2006; Park, Reder, & Dickison, 2005; Parks, & Yonelinas, 2007; Rotello, Macmillan, & Reeder, 2004; Tulving, 1985b; Wixted, 2007b; Wixted, & Stretch, 2004; Yonelinas, 2002). The remainder of this chapter is primarily concerned with the retrieval of information from long term memory (Tulving, 1983; see Figure 1-9) in simple item recognition, rather than associative or plurality recognition, which are thought to rely on different underlying processes (Westerman, 2001).

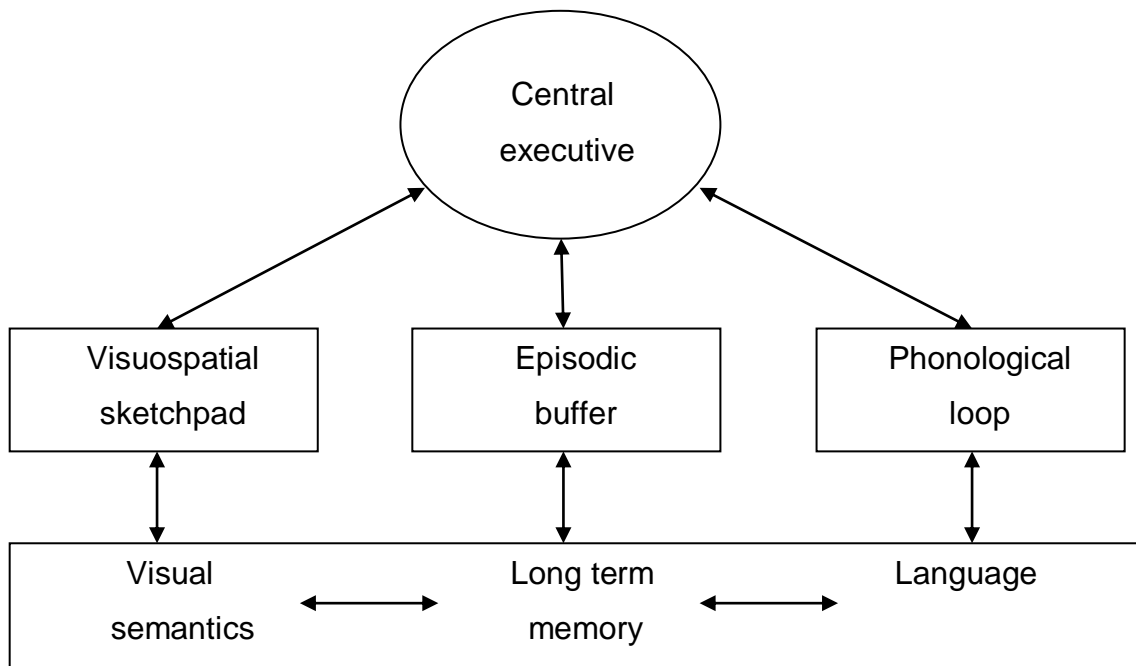


Figure 1-9: Human memory systems (from Baddeley, 2000).

Recognition is the awareness that something has been encountered before, and models of recognition memory are broadly divided into single- (e.g., Wixted, & Stretch, 2004) and dual-process models (e.g., Diana et al., 2006; Yonelinas, 2002). Current dual-process models of recognition memory (e.g., Atkinson, & Juola, 1973; 1974; Mandler, 1980; Tulving, 1985a; Yonelinas, 1994; 1997; 1999; Yonelinas, & Jacoby, 1996) assume that the recognition of previously encountered faces, objects or words occurs due to two independent mnemonic processes – recollection, a slow and effortful conscious process where specific contextual information concerning the original learning experience is retrieved from episodic memory, and familiarity, a relatively rapid and automatic process which provides a context-free sense that an item is known but without detailing why (see Yonelinas, 2002, for a review).

In contrast, single-process models (e.g., Donaldson, 1996; Gillund, & Shiffrin, 1984) propose that the qualitatively different experiences of recollection and familiarity originate from a single common neurocognitive process (Squire, Wixted, & Clark,

2007; Wixted, & Stretch, 2004). Therefore much of the debate in the literature centres upon whether recognition is based on one (familiarity) or two (familiarity and recollection) variables, whether the variable(s) draw on multiple sources of information (global, specific), and if a weighted summed source should be considered a single or multiple sources (e.g., Rotello et al., 2004, STREAK model) (Malmberg, 2008).

When recognition memory is explicitly tested under experiment conditions, participants are typically presented with a set of learning items (which may be written or auditorally presented words, images, faces, tones), followed immediately, or after a delay, by a second set of items containing the original learning items (old) and items that were not presented at learning (new), in the same modality or a different modality. There are different recognition tasks that may be employed, for example: 1) old/new task where participants just have to state whether a presented stimulus has been seen before in the experimental context; 2) rating task where participants also say how confident they are that they have (not) seen the stimulus before, according to a scale; 3) a two-alternative forced-choice task where participants have to decide which of two stimuli they saw before (see Malmberg, 2010).

Familiarity is proposed to be a continuous signal and stimuli such as words, pictures and everyday objects will already be associated with a certain amount of familiarity. However, a larger degree of familiarity is gained from exposure in the study phase of an experiment, allowing them to still be useful stimuli in standard old-new recognition tasks. When previously studied items (old) are intermixed with items not presented during learning (new), participants are typically able to correctly identify at least 70% of old stimuli (hits) and 80% of new stimuli (correct rejections) (Achilles, 1920; Yonelinas, 1994). More new items are correctly identified than old items, and there

are usually more misses (old items identified as new) than false alarms (new items identified as old) (Achilles, 1920). In contrast to familiarity, recollection is suggested to be an all-or-nothing threshold retrieval process, involving the recovery of episodic detail from the original learning instance (Yonelinas, 1994). Supposedly, in a recognition task, new items may evoke familiarity but they won't generate recollection, whereas old items elicit a stronger sense of familiarity, plus recollection, and hence result in a positive recognition decision.

Frequently, participants may be asked to decide whether they “remember” seeing an old stimulus at learning (conscious recollection leading to an R response), or whether they just “know” that it is old (feeling of familiarity leading to a K response), an introspective report known as a remember-know or R-K judgement (Gardiner, 1988; Tulving, 1985a). Source memory has been used as a more objective measure of whether the recognition decision is based on recollection or not, for example Yonelinas (1994; 2001b) asked participants which of two learned word lists each old item appeared on. If an item was identified as being on the correct list, and was given a high confidence rating, it was assumed to have been recollected. Researchers have used various terms to label the strength variables underlying recognition, such as global and specific (Rotello et al., 2004), item and associative (Murdock, 2006), and semantic and episodic (Reder et al., 2000), but this thesis will generally use familiarity and recollective strength (Wixted, & Stretch, 2004).

1.3.1.1. Single-Process Models

Widespread interest in recognition memory did not occur until the “cognitive revolution” in the 1960s because recognition was considered simpler and more straightforward than recall, partly due to its higher accuracy and perceived ease (Malmberg, 2010). Global memory models endeavour to account for task

performance in all conditions under a unitary theoretical framework (Malmberg, 2010). However, early models of recognition memory were signal detection measurement models (e.g., Banks, 1970; Bernbach, 1967; Kintsch, 1967; Lockhart, & Murdock, 1970), where recognition was based on the comparison of a single continuous variable, supposed to be memory strength or familiarity, to a criterion value. Judgement criteria are subjective, with conservative criteria generating fewer false alarms, but also fewer correct hits, and an individual's criterion may be modified by accumulating positive and negative feedback (cf., random walk theory; Ratcliff, 1978; Ratcliff, & Murdock, 1976; Murdock, 1985). Recognition occurs because old items are more familiar than new items, however, measurement models do not explain how stimulus familiarity is generated (Malmberg, 2008; 2010).

In the 1980s global matching process models were developed to explain the generation of familiarity in signal detection models of recognition (Clark, & Gronlund, 1996; Gillund, & Shiffrin, 1984; Hintzman, 1988; Humphreys, Bain, & Pike, 1989; Murdock, 1982; single-process recollection models are relatively rare, except Yonelinas, 1999; Diller, Nobel, & Shiffrin, 2001). Global matching models state that familiarity during a recognition test results from an assessment of the similarity between a stimulus and all information held in memory relating to the learning phase (Gillund, & Shiffrin, 1984; Hintzman, 1988; Humphreys et al., 1989; Murdock, 1982; Norman, & O'Reilly, 2003; Shiffrin, & Steyvers, 1997). The recognition decision is still based on a continuous variable, familiarity, which is generated by matching a retrieval cue (a transient representation of the stimulus) against the large number of traces in memory (Malmberg, 2008). The more similar a retrieval cue is to traces in memory, the more familiar it will 'feel', therefore as old items have been seen before in that context, and will closely resemble at least one trace, they will seem more familiar than new items (Malmberg, 2008).

Despite the benefits of being global models, which make few assumptions, global matching models in their present state were unable to account for phenomena such as list-length and -strength interference effects (Ratcliff, Clark, & Shiffrin, 1990), Receiver Operating Characteristic (ROC) curves (Ratcliff, Sheu, & Gronlund, 1992) and mirror effects (Glanzer, & Adams, 1985; Malmberg, 2008). In response to these challenges, new global matching models were developed with a Bayesian approach which assumes that memory systems have evolved to be optimal and adaptive, and aims to achieve maximal accuracy on the basis of the available information. For example, the Retrieving Effectively from Memory model (REM; Shiffrin, & Steyvers, 1997; 1998), the Theory Of Distributed Associative Memory model (TODAM; Murdock, 1997; 2006), the Subjective Likelihood Model (SLiM; McClelland, & Chappell, 1998) and the Bind-Cue-And-Decide Memory model (BCDMEM, Dennis, & Humphreys, 2001). Familiarity computations strengthen the likeness between retrieval cues and their trace, whilst lessening the likeness with other memory traces, known as differentiation (Criss, 2006).

Threshold models of recognition memory differ from signal-detection models (Krantz, 1969; Macmillan, & Creelman, 1990). The high-threshold model describes two item states, detected in memory and not detected in memory. A high threshold must be met for an item to achieve detection status, so only old items surpass the threshold, however false alarms can result from participants guessing when items are not detected (Malmberg, 2008). The double high-threshold model describes three item states and two high thresholds: a threshold only old items achieve to reach the detect-old state, a threshold only new items achieve to reach the detect-new state, and an indeterminate state for items that fail to reach either threshold, which may result in false alarms or misses (Malmberg, 2008). Threshold models are not widely accepted, partly because they predict the same manipulation effects on both single-

item recognition memory and recall, whereas many factors affect the two types of memory performance in different ways, for example: word-frequency (e.g., Balota, & Neely, 1980; Gregg, 1976; MacLeod, & Kampe, 1996), emotion (Hertel, & Parks, 2002), age (Craik, & McDowd, 1987), alcohol (Soderlund, Parker, Schwartz, & Tulving, 2005), primacy and recency (Achilles, 1920; Mulhall, 1915), and types of neurological impairment (Malmberg, 2010). These interactions have not been explained by threshold models; however, thresholds are often a component of dual-process models (Malmberg, 2008).

Whilst familiarity and recollection are *behaviourally* dissociable, it is unclear whether they are *neurally* distinct, i.e. whether separate anatomical structures or neuronal populations subserve familiarity and recollection (Rutishauser, Schuman, & Mamelak, 2008). Whilst several researchers propose that the hippocampus is concerned only with recollection (e.g., Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000; Holdstock et al., 2002; Yonelinas, 2001a), patients with hippocampal lesions often have a general loss of memory capacity rather than a specific recollection impairment (Manns, Hopkins, Reed, Kitchener, & Squire, 2003; Stark, Bayley, & Squire, 2002; Stark, & Squire, 2003; Wais, Wixted, Hopkins, & Squire, 2006). An fMRI study by Hannula and Ranganath (2009) showed that conscious recollection does not automatically accompany hippocampal activation during associative recognition with objects and scenes. Concomitant eye-tracking showed that participants fixated stimuli for longer on correct trials than incorrect trials, which also resulted in higher levels of activity within the hippocampus, but that this only lead to recollection when accompanied by prefrontal cortex activity (Hannula, & Ranganath, 2009).

Powerful evidence for single-process models of recognition memory comes from single-cell recordings. Rutishauser et al. (2008) made recordings from individual

neurons in the human hippocampus and amygdala, both part of the Medial Temporal Lobe (MTL), whilst epileptic participants performed an item recognition task. By asking participants to retrieve spatial locations of stimuli as well as their old/new status, they were able to determine whether they had been able to recollect episodic detail of the learning context. They found that neuronal activity increased in response to the second presentation (familiarity) of an old stimulus compared to the response to initial presentation at learning, regardless of successful recollection, but that the amount of change determined whether or not the stimulus location was recollected (Rutishauser et al., 2008). Consequently they concluded that human MTL neuron firing rates signal information pertaining to the phenomenological experiences of both recollection and familiarity, and proposed that their findings support a 'continuous strength of memory' model whereby stronger neuronal activity represents stronger memories (Rutishauser 2008; Rutishauser et al., 2008).

A different method of analysing EEG data involves looking at activity within different frequency oscillation bands (e.g., Klimesch, 1995). Gruber, Tsivilis, Giabbiconi and Muller (2008) analysed oscillatory EEG activity during a source discrimination recognition study of pictures of objects. They found that Induced Gamma Band Responses (iGBRs: 35-80Hz; 210-330ms) were not sensitive to source memory, whereas Induced Theta Band Responses were (iTBRs: 4.0-7.5Hz; 600-1200ms). iGBRs were higher for correctly identified "old" stimuli compared to "new" stimuli, suggesting that increased familiarity results from increased neuronal spike activity. Gruber et al. (2008) proposed that recollection was reflected in the theta band, whereas familiarity was reflected in the gamma band.

1.3.1.2. *Dual-Process Models*

Familiarity-only single-process models were criticised for being overly simplistic, and starting in the 1970s dual-process theories of recognition memory were developed, allowing recognition to be based on either item familiarity (signal detection), or episodic recollection of the learning context (threshold; best analogised by Mandler, 1980; Atkinson, & Juola, 1974; Kelley, & Wixted, 2001; Malmberg et al., 2004; Reder et al., 2000; Rotello et al., 2004).

The aim of dual-process theory is to quantify the recollective contribution to recognition (Malmberg, 2008). Many studies show recollection and familiarity can be differentiated behaviourally and these findings are used to support the argument that they have different underlying neural mechanisms (Yonelinas, 2002). The two types of response have been shown to respond differently to various manipulations (see Gardiner, & Java, 1993; Gardiner, & Richardson-Klavehn, 2000; Rajaram, & Roediger, 1997, for reviews). For example, “remember” judgements are impaired by the performance of a secondary task, such as auditory vigilance, whilst learning, whereas “know” responses are not (Gardiner, & Parkin, 1990). Repetition priming manipulations, where the stimulus is presented very briefly right before testing, enhance “know” responses but do not influence “remember” responses (Huber, Clark, Curran, & Winkielman, 2008). Changing the modality of stimuli between learning and test from pictures to words enhances “remember” responses but decreases “know” responses (Rajaram, 1993). Interestingly, manipulations which are known to affect remember judgements, also affect explicit tests of memory, and those known to affect know judgements also affect implicit tests of memory (Paller, Voss, & Boehm, 2007; Rajaram, & Roediger, 1997; Voss, & Paller, 2008; Yonelinas, 2002).

Dual-process models suggest that familiarity should act faster than recollection. Delayed recollection occurs in real-life as well as the lab, for example recognising someone's face but not knowing who they are until after they've walked past (Mandler, 2008; Mandler, & Boeck, 1974; Rabinowitz, & Graesser, 1976). Atkinson and Juola (1973) suggest that familiarity is activated initially as a fast search, with the slower, more thorough recollection process occurring only if familiarity is unsuccessful. Mandler's (1980) more recent "horse race" model proposes that both process occur in parallel, and that the relatively automatic familiarity process finishes before the more deliberate, intensive recollection search. It was Hintzman and Curran (1994) who first looked at this using a response-deadline procedure (Doshier, 1984; Gronlund, & Ratcliff, 1989; Hintzman, & Curran, 1997; Reed, 1973). At recognition, the time between stimulus presentation and participants' response was varied randomly. When participants were forced to respond quickly there was an increase in false alarm to similar lures, whereas when participants had longer to respond they made fewer false alarms. On average, participants were able to discriminate old and new words after 420ms, but took 520ms to discriminate old items from similar lures. The authors suggested that when distinguishing between old and new items familiarity was rapid and accurate, but increased the number of false alarms to lures, which were only correctly rejected once the slower recollection process failed to produce specific contextual detail (Hintzman, & Curran, 1994).

Hintzman and Curran have used a global-matching approach to behaviourally dissociate familiarity and recollection by manipulating the similarity between items at learning and recognition in a plurality recognition paradigm (Hintzman, & Curran, 1994; 1995; Hintzman, Curran, & Oppy, 1992). Participants were asked to remember plural and singular words (e.g., frog, books), including their grammatical number, and then were tested with old words (e.g., books), new words and plurality reversed lures

(e.g., frogs). Due to the higher familiarity of the lures, they produced many more false alarms than new words. By increasing the number of times items were presented during the learning phase up to 20, and asking participants to judge at recognition how many times the item had been presented (frequency judgements of new and similar words should be zero), they were able to manipulate familiarity (Hintzman, & Curran, 1995; Hintzman et al., 1992). Increasing presentation frequency increased frequency judgements (indexing familiarity of old and similar items), but not false alarms (indexing recollection of specific information about plurality; Hintzman, & Curran, 1995; Hintzman et al., 1992).

ERP studies provide evidence in support of dual-process models of recognition memory, due to the dissociable neural signatures evoked by recollection (late parietal) and familiarity (early mid-frontal; Curran, Tepe, & Piatt, 2006). Duarte, Ranganath, Winward, Hayward and Knight (2004) showed thirteen undergraduate participants 350 grey-scale pictures of common objects (e.g., duck, baseball) and asked them to judge either whether or not the object was alive, or whether or not it was hand-operated (source discrimination). At recognition 300 old stimuli were shown on screen for 180ms intermixed with 150 new stimuli, and participants were asked to make a recognition judgement, and if they judged it to be old, also make a R-K decision and an encoding category decision ("animate" or "manipulable"; Duarte et al., 2004). They found that items given a K response evoked an earlier positivity at frontal sites (150-450ms), and items given an R response evoked a positive-going ERP at frontal (300-600ms) and parietal (450-800ms) sites (recollection > familiarity > misses). Interestingly, ERPs recorded at encoding also differed between recollected and familiar items. Items later recognised on the basis of familiarity evoked a left-lateralised positivity at anterior sites (300-450ms), whereas items later recognised due to recollection evoked a right-lateralised positivity at anterior sites (300-450ms)

and bilaterally (450-600ms). Duarte et al. (2004) concluded that recollection and familiarity are manifestations of functionally, temporally and topographically dissociated patterns of neural activity during both encoding and retrieval. Numerous ERP studies have replicated these findings, showing very similar time windows, with familiarity occurring ~300-500ms and recollection ~500-800ms (see Düzel, Yonelinas, Mangun, Heinze, & Tulving, 1997; Wilding, & Rugg, 1997a; Rugg et al., 1998a; Curran, Schacter, Johnson, & Spinks, 2001; Curran, 2000; Wolk et al., 2006; Ally et al., 2008). Recollection and familiarity ERPs have also been doubly dissociated using manipulations such as picture superiority (Cohn, Moscovitch, & Davidson, 2010; Curran, & Doyle, 2011).

In contrast to neuropsychological evidence showing only general memory-impairment (Manns et al., 2003; Stark et al., 2002; Wais et al., 2006), or single cell recordings showing neurons which respond in situations of recollection and familiarity (Rutishauser et al., 2008), a double dissociation of recollection and familiarity exists whereby some amnesic patients with selective hippocampal damage have recollection impairments with intact familiarity (Aggleton, & Brown, 1999). Others have impaired familiarity with intact recollection, such as patient N.B. who had entorhinal and perirhinal cortex damage within the MTL (Bowles et al., 2007; 2010). Düzel, Vargha-Khadem, Heinze and Mishkin (2001) reported data from a patient with hippocampal damage who showed an absence of the Late Positive Component (LPC) in the 500-700ms window, normally associated with recollection, but preserved FN400 old/new effect in the 300-500ms window, thought to represent familiarity (see Chapter 6, section 6.1.1 for further description of these components). Such evidence may support dual-process models over unitary memory strength models (Squire et al., 2007), although recollection and familiarity may share a notable neuroanatomical overlap (Medina, 2008).

1.3.1.3. Evaluating Models

Responses can be viewed as occupying one of three regions of a two-dimensional ‘decision space’ (see Figure 1-10), which differ on whether a “remember” response depends on just recollection, the sum of recollection and familiarity, or the difference between recollection and familiarity (Rotello, & Macmillan, 2006). The process-pure (e.g., dual-process; Yonelinas, 2001b), model states that if there is sufficient recollective information, an R response will be made, and that a K response only occurs if there is sufficient familiarity but insufficient recollective information (Rotello, & Macmillan, 2006). If there is neither enough recollective or familiarity information, then a “new” response is made (see Figure 1-10A; Rotello, & Macmillan, 2006). In the one-dimensional (e.g., single-process; Wixted, & Stretch, 2004), model recollective and familiarity information is added together, so R responses are given to items for which the sum exceeds a higher threshold than for K responses, which are given to items for which the sum exceeds a higher threshold than for “new” responses, but lower than for R responses (see Figure 1-10B; Rotello, & Macmillan, 2006).

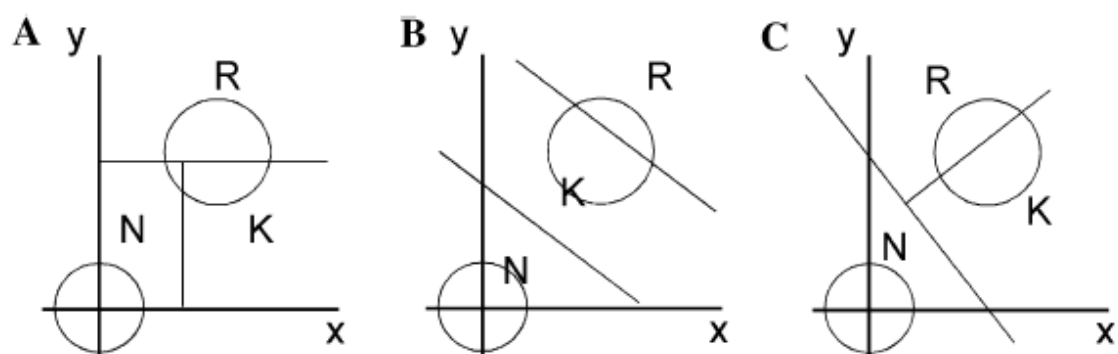


Figure 1-10: Decision space for the remember-know task (without ratings). (A) process-pure (dual-process) model, (B) one –dimensional (single-process) model, (C) sum-difference model (STREAK), x-axis shows familiarity/global memory strength, y-axis shows recollection/specific memory strength (from Rotello, & Macmillan, 2006).

In the STREAK model (Rotello, & Macmillan, 2006), either recollective or familiarity information means that an “old” response is given, but the strength components are oppositional, balancing out to establish the response (Rotello, & Macmillan, 2006). An R response occurs when there is relatively more specific contribution than global, and a K response occurs when there is relatively more global contribution than specific (see Figure 1-10C; Rotello, & Macmillan, 2006).

Rotello and Macmillan’s (2006) STREAK model is a single-process account. They asked participants to make binary and trinary R-K decisions with confidence ratings and found that 48/70 participants produced data that fit a one-dimensional strength model, and only 4/70 produced data that fit a dual-process model. They concluded that R and K responses depend on a single strength variable. By including confidence ratings to differentiate between models, Rotello and Macmillan (2006) may have changed the nature of the task into something that is more quantitative than the qualitative difference between recollection and familiarity. Also, recollective answers were excluded from the ratings scale (1-6, sure new to sure knew) for binary tasks, whereas new items were excluded from ratings (1-3 for details and feeling of knowing) for trinary tasks. In addition, they omitted the old-new paradigm, where the old-new decision is made first, then for old items an R-K judgement is made (see Rugg, & Yonelinas, 2003), which might fit a dual-process model.

However, although some data fit well, dual-process models have been criticised because it is difficult to separately and empirically estimate the contributions of recollection and familiarity; in addition recollection may simply be a stronger representation of familiarity, evoking additional detail. Single-process models are special instances of the more complex dual-process models, which revert to single-process when recollection does not occur (Malmberg, 2010). From the point of view

of philosophy of science, greater parsimony comes from single-process theories, which regard recognition as a strength continuum rather than separate categories (Curran, DeBuse, Woroch, & Hirshman, 2006; Medina, 2008), as it is not desirable to over complicate models when a single-process model is sufficient to explain observed data (see Diana et al., 2006; Dunn, 2004; Wixted, & Stretch, 2004; Yonelinas, 2002). A model should summarise the data with fewer parameters than data points, and not just repeat the data with a saturated model (Rotello, & Macmillan, 2006). Due to the lack of parsimony, some researchers have argued against the need for dual-process models, which also often don't explain how the memory signal is generated (e.g., Gillund, & Shiffrin, 1984). However, unlike single-process models, which are interested in memory strength, dual-process models are able to explain the dynamics and organisation of recognition memory, which single-process models are not (e.g., Atkinson, & Juola, 1974; Mandler, 1980).

Different models are better able to account for particular recognition memory phenomena, but this isolationist approach has failed to reach a wider consensus. Malmberg's (2008) framework explains more of the empirical data than other current models, including accuracy and retrieval dynamics of single-item recognition, associative recognition, and plurality discrimination. He proposes that an individual selects, from among several possible recognition strategies, the one most likely to generate an accurate answer most efficiently, thus accounting for the fit of different related models under different experimental conditions (Malmberg, 2008). This is supported by evidence showing that participants' decision rules vary depending on instructions. For example strategies are different when asked to make two consecutive binary decisions (whether an item is "remembered", before deciding whether non-remembered items are "known" or new), compared to when asked to

make a single trinary decision (“Is the item remembered, known or new?”) (Rotello, & Macmillan, 2006; Brown, & Bodner, 2011).

However, some researchers (e.g., Wixted, 2007a; Wixted, & Stretch, 2004; Rotello et al., 2004) propose that rather than being exclusive theories, dual-process and signal-detection can be integrated into a single model of recognition memory, such as the signal detection unequal variance model (Wixted, 2007a), STREAK (Rotello et al., 2004), and single-trace dual-process models (e.g., Greve, Donaldson, & van Rossum, 2010). These models assume that as well as familiarity, recollection also lies on a continuum, and that rather than recognition decisions being based on either recollection OR familiarity, both sources of memory information are combined into a unitary combined memory strength that is then compared with a criterion value to make a recognition decision (Wixted, & Stretch, 2004; Wixted, 2007a). This is supported by the fact that recollection can be graded, for example some contextual information recollected vs. all contextual information recollected (Ingram, Mickes, & Wixted, 2011; Wixted, 2007a). This view even unites apparently contrasting neuroanatomical evidence from lesion studies, and also that of the single-cell recordings made by Rutishauser et al. (2008). The hippocampal neurons measured as having increasing activity with increasing memory strength may be involved in summing the signal from separate populations of neurons representing familiarity and recollection.

It has been suggested that ERPs provide evidence in support of a combined memory strength model of recognition memory. Finnigan, Humphreys, Dennis and Geffen (2002) manipulated memory strength by repeating half of the old items three times during learning (strong) and the other half only once (weak), giving three ‘strengths’ of item at recognition – new, weak and strong. Finnigan et al. (2002) demonstrated that

the FN400 was sensitive to memory strength in a graded fashion, with new items being most negative, followed by weak and then strong items being most positive at parietal electrodes. The LPC was found to be sensitive to decisional factors such as confidence and accuracy, and upon visual inspection of the grand-averages appears to show the same pattern at parietal electrodes. They argue that their data provide evidence for a memory strength model.

It has been proposed that rather than distinguishing between recollection and familiarity, the traditional remember-know paradigm in fact distinguishes weak from strong memories (Wixted, & Mickes, 2010). Experimental design also influences participant strategy when deciding whether to respond “know” or “remember” (Kapucu, Macmillan, & Rotello, 2010; Rotello, & Macmillan, 2006). The issues of precisely how recollection and familiarity exist, what they comprise and how they act may still not be fully understood, but the recollection-familiarity distinction continues to be useful in studying recognition memory.

1.3.2. Psychophysiological Correlates of Recognition Memory Processes

The old/new paradigm is an informative experimental design that lends itself to combination with a psychophysiological technique, such as ERPs, fMRI, or pupillometry. Words or pictures learned during a study phase are mixed with new items, and participants respond “old” when they recognise an item from the study list and “new” to items that aren’t recognised. This paradigm can be used to measure differences in physiological responses to old and new items. The “old/new effect” was first established in the ERP literature, but has since been studied using ERPs, fMRI and Positron Emission Tomography (PET). Some of this literature is reviewed below.

1.3.2.1. *Event-Related Potentials*

ERP studies reveal more positive-going deflections associated with items correctly identified as old compared to items correctly identified as new, known as the ERP old/new effect (e.g., Karis, Fabiani, & Donchin, 1984; Maratos, Allan, & Rugg, 2000; Sanquist, Rohrbaugh, Syndulko, & Lindsley, 1980). There are two main ERP old/new effects reported in the literature, which consistently distinguish between old and new items during recognition memory tests, the parietal Late Positive Component (LPC) and the frontal N400 (FN400) old/new effects (for a review see Johnson, 1995). For example, Allan and Rugg (1997) demonstrated that in an old/new recognition test their 18 participants averaged more positive and longer-lasting left posterior ERPs for hits than for correct rejections. Whilst both ERP effects are present over both hemispheres, and larger for old items than new items, the frontal old/new effect, thought to reflect familiarity, is larger over the midline, and the parietal old/new effect, thought to reflect recollection, is larger over the left hemisphere (Allan, & Rugg, 1997; Curran, 2000; Curran, & Cleary, 2003; Curran et al., 2001; Curran et al., 2006; Goldmann et al., 2003; Rugg, & Allan, 2000; Vilberg, & Rugg, 2008; see Chapter 6, section 6.1.1 for further discussion of these components). The ERP old/new effect may provide a reliable, quantifiable marker of recognition processes, which can then be compared to participant report to link the experience of recognition to underlying neurocognitive activity.

The ERP old/new effect is comparable between epileptic patients who have and who have not had a medial temporal lobectomy (Rugg, Roberts, Potter, Pickles, & Nagy, 1991), and between participants with Alzheimer's Disease and healthy controls (Friedman, Hamberger, Stern, & Marder, 1992; Rugg et al., 1994). Verleger (1995) interprets this as meaning that the ERP old/new effect does not originate in the hippocampus. However, Smith and Halgren (1989) found that the ERP old/new effect

was reduced in participants who had had a temporal lobectomy. The variety of results is likely to reflect homogeneity in the precise lesions within these patient populations, and the effects of individual anatomy on the magnitude of the electric field measurable at the scalp (Luck, 2005).

1.3.2.2. *PET and fMRI*

Neuroimaging techniques like PET and fMRI are also able to detect distinct patterns of activity in response to old and new items. For example, a large number of fMRI studies have shown increased activity in the anterior left frontal cortex, and medial and lateral parietal cortex in response to correctly identified old compared to new items, suggesting involvement in item-related retrieval success (e.g., Browndyke et al., 2008; Donaldson, Petersen, Ollinger, & Buckner, 2001; Henson et al., 2005; Henson, Rugg, Shallice, Josephs, & Dolan, 1999; Konishi, Wheeler, Donaldson, & Buckner, 2000; Vilberg, & Rugg, 2008). Habib and LePage (1999) conducted a meta-analysis of PET studies and found increased blood flow to regions of the inferior and medial parietal lobe, and the left middle frontal gyrus, when participants viewed old items compared to new items. They concluded that stimuli need to be learned and tested in the same modality for this old/new effect to occur, suggesting a response to context/modality rather than just semantic or conceptual information.

1.3.2.3. *Pupil-Size*

A less studied marker responsive to memory processes is pupil dilation. As discussed in section 1.2, primarily since the 1960s, researchers have investigated the relationship between pupil-size and a variety of psychological processes (for reviews see Andreassi, 2000; Beatty, & Lucero-Wagoner, 2000; Goldwater, 1972; Janisse, 1977). Early studies did not specifically look at recognition memory, instead the majority of the research concentrated on “arousal” and “mental effort”.

Authors	Sample	Task & Stimuli	Measurement	Key Pupil-Size Findings
Beatty, & Kahneman, 1966		Digit recall: unfamiliar 7 digit number, and familiar name	Absolute, mm	10% increase following familiar name; larger pupil-size recalling well-known than unknown telephone number
Kuc, & Janisse, 1967		Digit span: digits		Larger on correct trials compared to incorrect trials
Craven, 1972		Words		Dilation to presented words, which could represent recognition
Gardner, Mo, & Borrego, 1974a	4 male under-graduate students	Recognition: nonsense words of consonant-vowel-consonant		Old > new
Gardner, Mo, & Krinsky, 1974b	4	Recognition: 3 consonant trigrams		Old > new
Stelmack, & Leckett, 1974	3 male graduate students	Recognition: taboo & neutral words	Standard photographic technique; mm	Decreasing pupil-size increases recognition threshold
Gardner, Beltramo, & Krinsky, 1975	6 college students	Delayed recall: digits	Craig video camera; measured from TV, 1Hz; grand average z scores of mean size	Largest during recall, increased during presentation, decreased during delay
Krinsky, & Gardner, 1977	6 under-graduate students	Comparison tone recognition: 1000Hz tone	Grand average z scores of mean dilation	Largest during recall, increased during presentation, decreased during delay
Gardner, Philp, & Radacy, 1978	7 children 7-9 yrs	Delayed recall: digits	Polymetrics pupillometer, manually scored from strip-chart; grand average z scores of mean size	Largest during recall, increased during presentation and decreased during delay
Garrett, Harrison, & Kelly, 1989	24 male under-graduate students	Passive viewing of pictures of nudes	Sony video camera, zoom f75 measured on high res monitor; log10 mean dilation	Slide x trial interaction – increased dilation on subsequent presentations for some slides
Maw, & Pomplun, 2004	20 under-graduate students	Recognition: pictures of famous/non-famous faces	SR Research EyeLink II; PDR of trial pixel area to baseline at initial setup	Famous faces > non-famous faces
Otero, Weekes, & Hutton, 2006	36 under- & post-graduate students	Old/new, R-K DRM recognition: words & pictures	SR Research EyeLink II & Data Viewer; PDR of max size during trial to participant average across trials	Old > critical distracters > new items; remember > know

Authors	Sample	Task & Stimuli	Measurement	Key Pupil-Size Findings
Laeng et al., 2007	3 amnesic patients	Old/new recognition & confidence: colour pictures	SensoMotoric Remote Eye-Tracker; single baseline to blank white screen from single patient	New > old
Võ et al., 2008	19 students	Speeded old/new recognition: positive/neutral/negative words	SensoMotoric Hi-Speed eye-tracker, 250Hz; peak dilation minus baseline	Old > new; old neutral > old positive > old negative > new negative > new positive > new neutral
Otero, Weekes, & Hutton, 2011	45 students	Old/new, R-K recognition: picture names	SR Research EyeLink II & Data Viewer; PDR of max size during trial baseline	Remember > know > new
	34 under-graduate	LOP, old/new, R-K recognition: spoken nouns		Deep > shallow > new
	37	DRM recognition: object names		Old > false alarms > new
Heaver, & Hutton, 2011	26	Old/new recognition: words, feigning memory loss	SR Research EyeLink II & Data Viewer; 500Hz, PDR of max size during trial to trial baseline	Old > new under standard instructions and when asked to malingering or respond “new” to all items
Kafkas, & Montaldi, 2011	41 psychology under-graduate	Old/new, R-K encoding: object pictures	Eye-Trac 6000; 60Hz; peak pupil diameter in mm minus average trial baseline size	At encoding misses>familiar >recollected
Papesh, & Goldinger, 2011	30	Old/new; auditory low/high frequency words	Tobii 1750; 50Hz, average pupil diameter and baseline diameter	Old>misses/new particularly for low frequency words
Papesh, Goldinger, & Hout, 2011	29 students	Old/new recognition: spoken words & nonwords in two voices	Tobii 1750, E-Prime; 50Hz, peak diameter minus average diameter during trial baseline	At encoding hits>misses; at encoding & retrieval nonwords>words; at retrieval original>familiar> new voices
Naber, Rutishauser, & Einhäuser, unpublished	32	Old/new recognition: photographs; confidence	SR Research EyeLink 2000 & Matlab, 500Hz; diameter, no baseline	More constriction to new items than old items
Van Rijn, Dalenberg, Borst, & Sprenger, submitted	19 psychology under-graduate	Learned names of brain areas	SR Research EyeLink 1000; 500Hz; percentage change relative to baseline	Pupil response decreased with repeated presentations

Table 2: Studies of recognition memory and pupil-size; PDR = Pupil Dilation Ratio, see section 2.1.2.1.

However, some studies reported results that could also be interpreted as demonstrating recognition effects of pupil-size (see Table 2). For example, Craven (1972, cited by Janisse, 1977) observed dilation to presented word stimuli. Also in Kuc and Janisse's (1967, cited by Janisse, 1977) digit span study, the larger pupil-size on correct trials compared to incorrect trials could represent a stronger memory signal leading to recall success.

In recent years studies have begun to look directly at recognition memory and suggest a possible relationship with pupil-size. In probably the first LTM pupil study, Beatty and Kahneman (1966) investigated pupil-size and memory load, finding the same sort of pupil-size changes to processing load as occurs with digit recall in STM tasks. They compared pupil responses when participants recalled an unfamiliar 7 digit number provided by the experimenter, compared to recalling their own telephone number from long-term memory. Long-term memory retrieval of a well-known telephone number evoked larger pupil-sizes (0.5mm) than the seemingly more effortful recall of an unknown number (0.34mm). They suggested that the pupil reflected the retrieval of information from long-term memory (Janisse, 1977). Gardner, Mo and Borrego (1974a) presented four participants with previously seen ("well-formed" memories created during the learning phase) and unseen (not presented during learning) nonsense words comprising a Consonant, Vowel and Consonant (CVC). Gardner et al. (1974a) reported pupil dilation to "old" CVCs, and constriction to "new" CVCs for all four participants. Following on from this study, Gardner, Mo and Krinsky (1974b) attempted to replicate the results using high frequency words presented in the auditory rather than visual modality to guard against pupil-size changes as a result of visual features (see section 1.2.1.1). This time they found no significant differences between pupil dilation to old and new words. However, their sample size was underpowered with again only four

participants, and Gardner et al. (1974b) concluded that stimuli were too high frequency and were equally familiar regardless of whether they were on the learning list or not. However when using strings of three randomly generated consonants (trigrams) in their second experiment, to remove pre-existing stimulus familiarity, they found that mean pupil-size increased more to old learned items than to novel items (Gardner et al., 1974b). Gardner, Beltramo and Krinsky (1975) felt pupil-size reflected the cognitive effort arising from the storage and retrieval from memory, and found constriction during retention when participants reported rehearsing information. Gardner et al. (1978) suggested that rather than indicating general mental effort, “pupillary dilation is specific to mental encoding and retrieval of information” (p. 168).

More recently Maw and Pomplun (2004) showed 20 undergraduate participants 40 famous and 40 non-famous faces, with equal numbers of male and female faces. Wearing an EyeLink II eye-tracker, participants were asked to press a button to indicate whether or not they recognised each face. Although the main focus of the study was the eye-tracking data, they found that maximum pupil-size increased relative to baseline in response to famous faces but not to non-famous faces (Maw, & Pomplun, 2004). The authors asserted that pupil-size represents memory processes associated with recognising a face, however they did not test pupil responses to non-famous familiar faces, or non-face stimuli (Maw, & Pomplun, 2004).

In the first robust study using modern eye-tracking methods to look explicitly at recognition memory, Otero, Weekes and Hutton (2006) showed 36 participants words and pictures of everyday objects, and found that maximum pupil-size was consistently larger when participants viewed old items previously encountered during learning, compared to new items, independent of encoding modality (pictures vs. words). In addition, pupil-size in response to semantically-related lures was also larger than for

correctly identified new items. In two follow-up studies the authors replicated their findings using concrete nouns presented visually, and extended this to show that the Pupil Old/New Effect (PONE) also occurred with spoken word stimuli (Otero, Weekes, & Hutton, 2011).

Other researchers have confirmed the PONE. Võ and colleagues (2008) showed 19 participants words varying in emotional content (positive, negative and neutral) to investigate the influence of affect on pupil-size during word recognition. During a rapid recognition test they found larger pupil-sizes to correctly classified old words than correctly classified new words, and that the PONE was reduced for words with positive or negative emotional valence. They claimed to have introduced the PONE for the first time, however, as discussed above this is not strictly the case, although the emotional attenuation of the PONE was a novel effect. An alternative explanation for some findings may be that stored information, such as a face or telephone number, has emotional associations that enlarge the pupil (see section 1.2.2; Janisse, 1977). Porter et al. (2007) state that cortical areas active during tasks evoking pupil responses are closely interconnected with areas implicated in memory, such as the limbic and reticular activating systems, which are also involved in emotional arousal (Brown et al., 1999; Löwenstein, & Loewenfeld, 1962). Silk et al. (2009) found that recall of emotional words evoked larger pupil-sizes than non-emotional words, however Võ et al.'s (2008) recognition paradigm supports earlier findings of a U-shaped pupil-size function where larger pupils are found in response to neutral stimuli than for slightly positive and negative stimuli (e.g., Levine, & Hakerem, 1969, cited in Janisse, 1974; Gunther, & Lussier, 1975, cited in Janisse, 1977).

In explaining their findings, Võ et al. (2008) proposed that the PONE represents the greater cognitive effort required to correctly identify old compared to new stimuli,

based on extensive previous research demonstrating the relationship between pupil-size and cognitive effort (see section 1.2.2.1). They argued that recollection requires the retrieval of qualitative contextual information, including the experience of an old item during the study phase, which is more cognitively demanding than the correct rejection of a new item, which does not. Võ et al. (2008) suggested that the attenuated pupil response to emotionally valent words reflects the relative ease, and therefore reduced cognitive load, with which words are recognized due to their associations.

Whilst building on a substantial body of research demonstrating links between pupil-size and cognitive load, there are problems with a cognitive load account of the PONE. Firstly, although the correct rejection of new items may not involve precisely the same recollective processes as occur during recognition of old items, it is not clear why recognition of previously presented items should necessarily be more cognitively demanding than the correct rejection of novel items, and it is certainly not the case that *no* cognitive effort is involved. Correct rejection may involve an effortful memory search, and studies have found that it typically takes longer than correct recognition (e.g., Ratcliff, & Murdock, 1976), particularly for items involving “recall-to-reject” (Leding, & Lampinen, 2009). For example, in a remember/know recognition memory ERP paradigm, Wiese and Daum (2006) found that the average response time for hits was 1144ms, whereas the average response time for correct rejection of non-critical lures was 1355ms. This suggests that recognizing an old item is not necessarily more cognitively demanding than correctly rejecting a new item.

An alternative interpretation is put forward by Otero et al. (2011) who advance that, like Finnigan et al.’s (2002) graded memory strength ERPs, the PONE represents a combined memory signal strength. They suggest that recognition of old stimuli is

more automatic than rejecting new stimuli, which may be somewhat familiar but generate no further detail on which to base a decision (Otero et al., 2011). Their proposal is supported by findings that pupil-size is larger for items which are recollected compared to items which are known, and is larger for known than new items. In addition, pupil-size for false alarms (new items incorrectly identified as old) is intermediate between correctly identified old and new items. Otero et al. (2011) argue that both familiarity and recollection vary on a strength continuum, and that old stimuli elicit a stronger summed familiarity and recollection signal than new stimuli, leading to larger pupil-sizes as a direct result of the greater combined memory strength. This explanation is supported by Papesch, Goldinger and Hout (2011) who found that “stronger” memories were associated with larger pupil-sizes than weaker memories. Another interesting finding was that items later correctly recognised evoked larger pupil dilations at learning than items that were subsequently forgotten, an effect also demonstrated in the ERP literature (e.g., Karis et al., 1984; Uhl et al., 1990; Fabiani, & Donchin, 1995), and suggesting greater effort went into encoding (Papesch et al., 2011).

1.4. Summary

Beatty and Lucero-Wagoner (2000) sum up by saying, “Pupillometry has served psychophysiology well in the study of the dynamics of human cognitive processing” (p. 159). Pupillometry is one of the more affordable psychophysiological techniques, is portable, non-invasive, and does not rely on behavioural responses. Changes in pupil-size consistently and reliably report the time-course of within-task, between-task and individual variations in cognitive processing, so the relatively compact literature is surprising. Despite an entire chapter on pupillometry in the second edition of the *Handbook of Psychophysiology*, by the third edition in 2007, pupillometry is not mentioned. Pupillometry may still be trying to dispense with the bad reputation some

researchers gave it in the 1960s, but equipment and techniques have improved and there has been a recent revival in its application to areas such as cognitive load, emotion processing, deception, and recognition memory.

The following chapters and series of experiments explore the PONE under a variety of conditions. Experiments 1 and 2 aim to replicate the PONE in a standard explicit test of memory, and determine whether a similar effect can be observed in an “implicit” test of memory. It is clear that the PONE is now well established for “explicit” recognition, but as yet it is still not clear what exactly the effect represents, whether it is associated with specific a mnemonic process, or whether an old/new effect can also be observed when memory is tested “implicitly”.

Experiments 3 and 4 aim to further investigate the mnemonic processes associated with pupil dilation by measuring pupil-size in an Artificial Grammar Learning (AGL) condition, proposing that for the implicit condition, conscious recollection will not be available, as both the “grammatical” and “nongrammatical” strings presented in the recognition phase will be different to those presented in the learning phase. Previous research (Reber, 1967; 1969; Scott, & Dienes, 2008; 2009) has indicated that implicit learning of grammatical letter strings evokes a greater sense of familiarity than non-grammatical strings. This would suggest that familiarity alone is sufficient basis for a recognition judgement. It was predicted that a PONE would occur in the implicit condition, reflecting familiarity signal strength, but that this effect would be smaller than that in a standard test of memory.

Experiments 5, 6 and 7 use changes in pupil-size to explore the role of conscious awareness in the PONE by drawing on the psychophysiology of deception and malingering literature. Experiment 5 explores whether the PONE is under voluntary control by asking participants to perform at their best, to deliberately perform poorly,

or to respond “new” to all items, in a standard recognition memory test. If, like the ERP old/new effect, the PONE is not under voluntary control, pupil-size should increase for old items compared to new items, even when participants say that they do not recognise stimuli. Next Experiment 6 provides participants with instructions for three different types of malingering strategy that might be used – not paying attention during learning, randomly preloading a response, and not responding during recognition. This experiment aims to artificially reduce performance measures during a standard test of recognition memory and observe any effects on the PONE. Then Experiment 7 asks participants to perform a secondary task during learning and recognition in a divided attention paradigm, and looks at how genuinely reduced recognition performance (simulating memory-impairment) affects the PONE. It was predicted that interfering with the encoding and/or retrieval of stimuli would reduce the magnitude of the PONE compared to when participants performed a single task at learning and recognition.

The last empirical chapter contains Experiments 8 and 9, which explore the effect of a graded memory strength manipulation on the PONE, in line with Otero et al.’s (2011) memory strength explanation, and explores the idea that the old/new effects seen in ERPs and pupil-size may index the same mnemonic processes. Few studies have measured ERPs and pupil-size simultaneously, with none having examined recognition memory specifically, therefore Experiment 8 recorded concurrent ERP and pupil-size data. It was predicted that the strength manipulation would also produce a graded effect whereby pupil-size was larger for strong items (seen three times at learning) than weak items (seen once at learning), and larger for weak than new items due to the differences in memory strength.

2. Methods

In Chapter 1, the pupillometry literature and relevant recognition memory literature was reviewed. The aim of Chapter 2 is to describe the general methods used to collect data for this thesis.

2.1. Pupillometry

2.1.1. Background

Like fMRI, ERPs and other psychophysiological measures, great care must be taken when attempting to draw inferences about cognitive function from pupil-size data. One significant problem pertaining to the interpretation of any continually changing signal (such as changes in pupil-size) is identifying individual contributions. For example, during any task there may be several (potentially overlapping) events that cause changes in pupil-size. Whether or not the contribution of individual events can be quantified depends on careful experimental design and use of control conditions where the only variable thought to change between conditions is the one of interest (e.g., Partala, & Surakka, 2003; Oliveira et al., 2009).

2.1.1.1. Techniques

One of the earliest references to a pupillometer is by Archimedes (212-187 BCE, cited by Schweitzer, 1956). However, it was not until the nineteenth century that more objective photographic methods were developed (for a fascinating review of the history of pupillometry see Hakerem, 1967). At the time of the resurgence of interest in pupillometry during the 1960s, most studies made use of 16mm movie cameras with mirrors and lenses to enlarge the eye. Images were recorded on infra-red film at 1-2Hz, and once the film was developed vertical or horizontal pupil diameter was

measured manually with a ruler or grid from an average of 20 individual frames per stimulus, which were projected onto a screen or table (e.g., Hess, & Polt, 1960; 1964; Kahneman, & Beatty, 1967; Kahneman et al., 1969; Paivio, & Simpson, 1966). For known camera-to-eye distances and magnifications, actual pupil-size in mm could then be calculated from measured pupil-size (Janisse, 1977). Popularised by Hess (1965), infra-red photography had the advantage that infra-red light does not trigger the light reflex, is reflected by the iris whilst being absorbed by the pupil, and is still detectable in a variety of light conditions (Hakerem, 1967).

Hand measurement was time-consuming and imprecise; Janisse (1977) reported a colleague manually measuring 100,000 frames for a single study. The first device that measured changes in pupil-size “online” electronically with a signal processor was the 60Hz Löwenstein Pupillograph developed by Löwenstein and Loewenfeld (1958) which scanned the eyes with a low intensity infra-red beam (Hakerem, 1967). This type of photoelectric device was developed into “television” pupillometers during the 1970s and over the intervening years the technology, resolution ($\sim 0.001\text{mm}$) and sampling rates (up to 1000Hz) vastly improved. Modern pupillometry research uses mobile or semi-mobile video-based eye-trackers, such as the EyeLink II (SR Research, Ontario, Canada), which are often head-mounted with two small infra-red cameras angled towards the eyes, and measure variables such as gaze position, saccades, blinks, fixations and pupil-size (see Figure 2-1a) (Wang, 2010).

A key advantage of modern pupillometry is that it is non-invasive; in particular remote infra-red eye-tracking equipment, such as the EyeLink 1000 (SR Research, Ontario, Canada), can be used to measure pupil-size from a desktop position without the need for head-mounted equipment or restraints such as chin and head rests (see Figure 2-1b). This is of particular benefit with populations or paradigms where a head-

mounted or restrained eye-tracker would interfere with the task or be impractical, for example studies involving developmental populations (e.g., Chatham, Frank, & Munakata, 2009) or concurrent ERP acquisition.

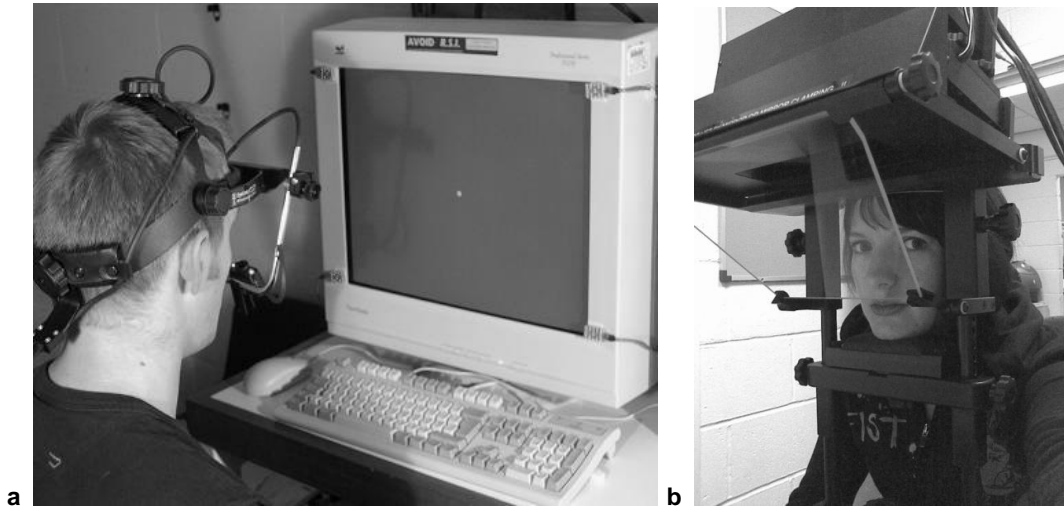


Figure 2-1: (a) Head-mounted and (b) tower-mounted EyeLink eye-trackers.

Yet, in remote set-ups the camera is further from the eye (50-100cm) giving less precise measurements and meaning that the pupil-camera distance has to be estimated for each frame, due to unrestrained head movement (Klingner, 2010). As very few pupillometry studies have been conducted using remote eye-trackers (e.g., Klingner, 2010; Klingner, Kumar, & Hanrahan, 2008), these systems are less validated and findings less replicated than those with head-mounted eye-trackers. However, researchers are developing minimal-calibration and calibration-free eye-tracking techniques (Hansen, & Pece, 2005; Ohno, & Mukawa, 2004), combined with increasing affordability and availability, this situation will soon change.

All experiments reported in the present thesis measured pupil size (and gaze position) using either a head-mounted EyeLink II eye-tracker (Experiments 1, 3, 4, 5, 6, 7 and 9) or a tower/desk-mounted EyeLink 1000 (Experiments 2 and 6), both manufactured by SR Research, Ontario, Canada. As gaze-tracking requires precise

localisation of the centre of the pupil, the EyeLink eye-trackers routinely and precisely measure pupil-size in camera pixels as a by-product. When not fixating centrally, the pupil becomes distorted (an ellipse), and whilst pupil size can be approximated by the eye-tracker using a foreshortening division, this links pupil size with gaze position and introduces a potential confound. Distortion also means that the pupil area measure is more stable than pupil diameter, which is calculated based on the assumption that the pupil at a particular moment is circular, and which will not be the case if the participant is looking away from the centre of the screen. Therefore, in the experiments reported here, small stimuli were presented in the centre of the screen and participants were asked to look straight ahead.

2.1.1.2. Data Acquisition

Once participants were seated comfortably with the head-mounted eye-tracker, or with their chin in the chin-rest of the tower-mounted eye-tracker, a nine-point calibration and validation procedure was carried out to ensure test-retest accuracy of $<0.5^\circ$ of visual angle. Whilst this procedure is less critical for pupillometry studies than gaze-tracking studies, a good calibration can only be performed if the thresholds for pupil colouring have been set properly. An automatic thresholding option was used to set the pupil colouring threshold – the greyscale threshold at which the host PC determines that a dark circular area is the pupil (see Figure 2-2).

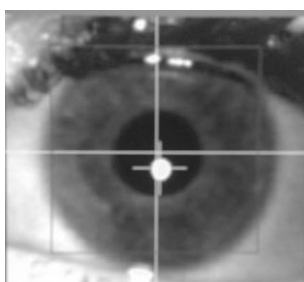


Figure 2-2: Pupil (shown as blue on computer display) as located by eye-tracking software.

2.1.2. Pupil-Size Reporting Variables

There are many different ways to report the effects of task on pupil-size, with a number of interconnected issues: should measurements report a size metric (minimum, average, maximum or difference) or latency (onset, offset or peak)? If size, should it be in diameter or area, measured in absolute units (mm) or relative units (percentage or ratio)? Should results be adjusted with a pre- or post-stimulus baseline, or an overall average (Janisse, 1977)? For example, is a 1mm or a 10% change to a small pupil equivalent to a 1mm or 10% change when the pupil is already large? An apparently equivalent change in diameter is very different when considering the change in area, with larger increases for pupils with larger baselines. What if two participants end up with a final pupil-size of 7mm diameter, but started from different baselines – is it fair to conclude that one participant made more effort (see Chapter 1, section 1.2.3), that the other was more anxious/motivated/aroused (see Chapter 1, section 1.2.2), or that this task may have a maximum processing load generating a maximum pupil-size? To explain the approach used in this thesis, these issues are first given further consideration.

Wilder's (1958) Law of Initial Values (LIV) states that "the change of any function of an organism due to a stimulus depends, to a large degree, on the prestimulus level of that function" (p. 199), meaning that trials with a larger baseline pupil-size will show a smaller increase in response to the stimulus than trials with a smaller baseline pupil-size, and vice versa. The smaller pupil has more "room" to change, whereas the larger pupil may experience a "ceiling" effect.

According to Janisse (1977), by the mid 1970s 90% of pupillometry studies reported relative changes in diameter as a percentage of a baseline, with only a small number reporting percent change in pupil area. Later studies commonly reported pupil

diameter change in millimetres by subtracting the baseline from the trial peak, which Beatty and Lucero-Wagoner (2000) suggest is a more complete and appropriate measure. They also argue that stimulus-evoked pupil-size changes have been demonstrated to be independent of baseline diameter across a wide range of initial values, tasks and laboratories (e.g., Bradshaw, 1969; 1970; Beatty, 1982b), and feel that smaller baseline values inflate percentage measures of pupil-size change (Beatty, & Lucero-Wagoner, 2000).

Dureman and Scholander (1962) highlight the antagonistic nature of the psychosensory dilation and light reflexes, whereby as the pupil dilates in response to a stimulus, the additional light falling on the retina triggers constriction of the sphincter muscle in opposition to the dilator muscle. These influences are not necessarily linearly related, and Dureman and Scholander (1962) suggest that because resistance from the sphincter increases “as a positive function of the... pre-stimulus pupillary area” (p. 51), it generates more opposition when the pupil dilates from 5.5 to 6.5mm, than from 3.0 to 4.0mm. The absolute change in diameter is the same, whereas the pupillary area changes by 12mm^2 and 7mm^2 respectively, reflecting the larger amount of activity required to produce a 1.0mm change in an already larger pupil. They therefore prefer area measures for both changes in pupil-size and maximum dilation (Dureman, & Scholander, 1962).

Janisse (1977) suggests that the “best” pupil-size index may be context-specific, and that no single measure is suitable for all experimental situations. For modern eye-trackers, calibration errors, and individual differences such as eye size, camera-pupil distance, and the refractive power of the cornea and participant glass/contacts, lead to difficulties in back-calculating absolute size from camera pixel-count. Recent research has therefore been carried out using relative percentage and ratio measures

of area rather than absolute measures of diameter (e.g., Bailey, & Iqbal, 2008; Heaver, & Hutton, 2010; 2011; Hupé et al., 2009; Maw, & Pomplun, 2003; Kang et al., 2009) and relative measures more easily allow for comparisons between individuals and groups, as they account for individual differences in baseline or peak through normalisation (e.g., Conati, & Merten, 2007).

There is no statistical test to directly compare experimental effects within a group to the same effect within a different group (manipulation by group interactions in a between-group design) due to pre-existing differences in the variable of interest, or level of noise, leading to non-equivalent groups (Luck, 2010; Nieuwenhuis, Forstmann, & Wagenmakers, 2011b). One way in which researchers have attempted to address this issue is to compare relative rather than absolute effects. For example, if in an experiment a group of participants in condition A have an average pupil-size measured by the eye-tracker as 2,000 camera pixels larger for old items than new items, but another group in condition B only show a difference of 200 pixels larger for old then new items, this would produce a significant main effect of group and a significant item-type by group interaction based on absolute values. If group A's pupils were larger to start with, and in fact changed from 8,000 to 10,000 pixels, whereas group B's pupils were smaller and changed from 800-1000 pixels, this is a relative change of 25% for both groups, and an analysis would reveal a significant main effect of item-type, but no interaction with condition and no main effect of group (see Figure 2-3). Some experiments require a between-group design, for example if naïve participants are needed in both conditions, therefore a relative measure of change helps to counter between-group differences.

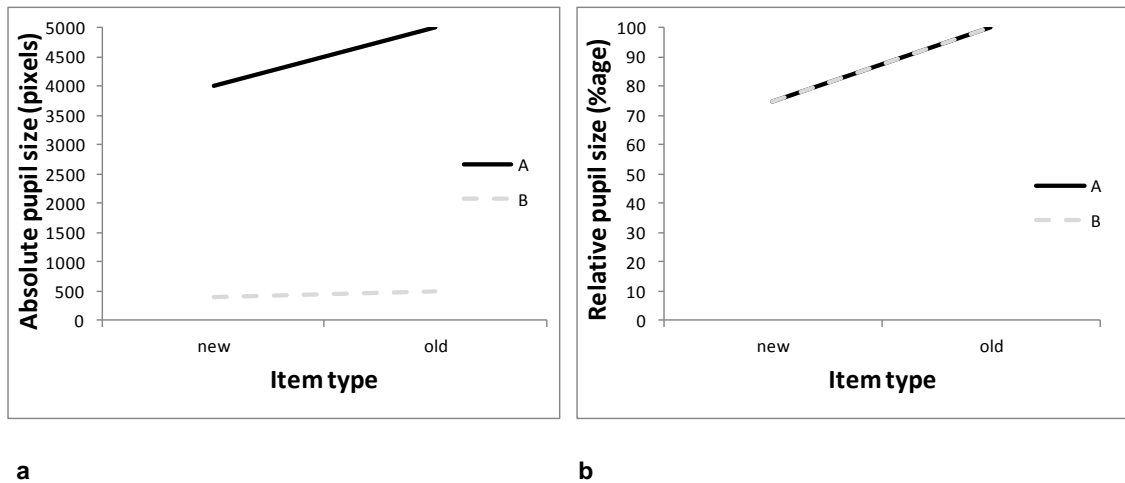


Figure 2-3: (a) Experiment reported as absolute pupil-size values, (b) Same experiment reported in relative values.

Having considered the issues above, the output variables of our equipment and the current standard practise with modern eye-trackers, we decided to use an area measure rather than a diameter measure, and a relative measure (pixel ratio) rather than an absolute measure. As this thesis is concerned with the magnitude of the memory signal we chose to use a size change metric rather than latency. Maximum pupil-size was used rather than average pupil-size because although maximum measures may be sensitive to random noise at the peak, making the maximum slightly larger than the true value, the average measure would mean excluding a large proportion of trials where participants either blinked, looked around the screen or the eye-tracker momentarily lost the pupil (situations which reduce measured pupil-size).

2.1.2.1. Pupil Dilation Ratio

The EyeLink II (500Hz) and EyeLink 1000 (1000Hz) eye-trackers used to collect pupil-size data for this thesis provide an arbitrary unit of measurement, reflecting the number of camera pixels occluded by the pupil image as determined by the EyeLink host software, together with other metrics including number and duration of fixations, eye position and blinks, which can be analysed in Data Viewer (SR Research,

Ontario, Canada). The number of pixels occluded by the pupil typically falls between 800-2000 units (± 1 unit) and 10% of variance is due to factors such as the distance between the camera and the eye, angle of the camera (the EyeLink II camera is positioned below the eye rather than in front of it where it would occlude vision; the EyeLink 1000 when positioned on the desktop is below eye-line and the angle is therefore affected by participant height), gaze position, and individual differences such as pupil position, corneal distortion, resting pupil-size and overall eye size. The measurements are difficult to convert to absolute units, and whilst diameter is measured, area is recommended by the manufacturers (SR Research, Ontario, Canada).

In order to gauge the degree of stimulus-evoked pupil response and generate a comparable measure, Maw and Pomplun (2004) devised a Pupil Dilation Ratio (PDR) by dividing maximum trial pupil-size by a single baseline measured immediately after initial calibration of the EyeLink II, and found PDR was significantly larger to famous faces than non-famous faces. However, because during long experiments the iris muscle fatigues (Löwenstein, & Loewenfeld, 1964; Peavler, 1974), stimulus-evoked responses diminish (Francis, & Kelly, 1969; Lehr, & Bergum, 1966; Löwenstein, & Loewenfeld, 1952) and baseline pupil-size decreases due to autonomic arousal decrement (Lehr, & Bergum, 1966; Sternbach, 1966; Woodmansee, 1966), the experiments in this thesis took a baseline measure at the beginning of each trial (Otero et al., 2011). PDRs reported here represent the maximum pupil-size during the 1750 or 2000ms trial period as a proportion of the maximum pupil-size during the 250 or 200ms pre-stimulus baseline period (Otero et al., 2011).

To reduce fatigue, and the potential effects of loss of interest or boredom, experiments were also limited to 30 minutes of measurement, stimuli were presented

in a random order, and rest breaks were offered between blocks of more than 50 trials (Klingner, 2010; Sternbach, 1966). In an experimental trial, pupil-size usually reaches maximum >1000ms after stimulus presentation (Beatty, 1982b), whereas after ~2000ms participants may lose focus on the stimuli, and occasionally look away from the centre of the screen, which can lead to confounds in pupil-size data (Otero, 2010). Therefore recognition trials were 2000ms long in order to ensure maximum dilation was captured. Data was recorded from one eye (typically the right eye) because the pupils are yoked (see Chapter 1, section 1.1.1; Reeves, & Swenson, 2004). Mean pupil-size can offer a more robust measure of response in situations where trials differ in length, however trials in this thesis are of equal length between participants, therefore measurements of maximum pupil-size were recorded rather than mean pupil-size, consistent with the literature.

2.1.2.2. Measurement Issues

As argued earlier, not all changes in pupil-size are necessarily due to the experimental effect under investigation, and Loewenfeld (1958) reports that externally triggered changes in pupil-size are overlaid on a signal with a variable level of noise. It is therefore highly probable that changes in pupil size caused by “internal” events are also superimposed on this varying signal (see Figure 2-4). As discussed in Chapter 1 section 1.1.2.5, one source of background noise is endogenous pupillary unrest, or hippus (Woodmansee, 1966), which may change diameter by 1% every second, and up to 10-20% every few seconds (Woodmansee, 1966). Hippus is amplified by fatigue and passivity, and suppressed by alertness and mental activity (Bouma, & Baghuis, 1971; Kahneman, 1971; Miller, & Newman, 2005). This means that a pre-stimulus baseline measure of pupil-size may include more hippus than the trial measurement. Researchers have taken a variety of approaches to dealing with hippus, including averaging over repeated measures of at least 8 trials per participant

to increase the signal-to-noise ratio (Hakerem, & Sutton, 1964; Woodmansee, 1966), selecting participants who are familiar with the testing environment and procedure, and who are alert and well-rested (Janisse, 1977), and using range correction, designed for assessing heart-rate and electrodermal responses, to reduce noise by computing each individual's possible range of pupil-sizes and expressing the actual value as a proportion of the individualised range (Lykken, 1972).

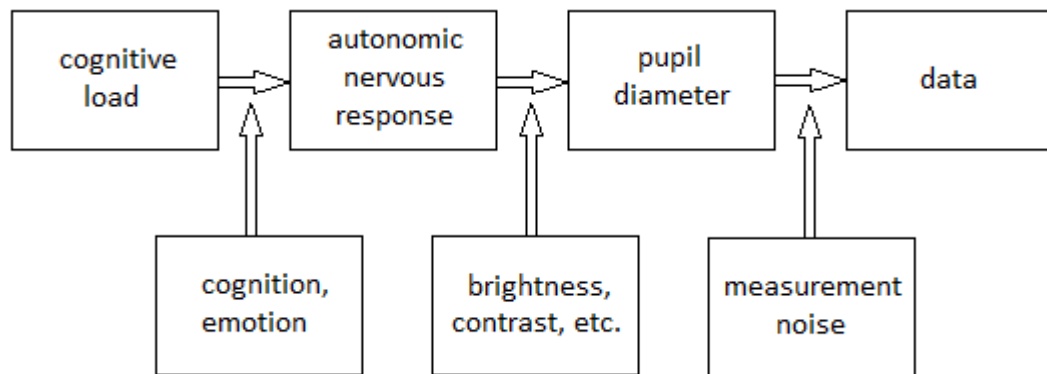


Figure 2-4: Sources of variation in measurements of pupil diameter (from Klingner, 2010).

However, Kahneman (1973) was confident that task-related focus was sufficient to reduce pupillary hippus, stating that changes in pupil-size are so reliable and predictable, that he took no further steps to control for it. In order to minimise the influence of artefacts, the baseline measure of pupil-size in this thesis is maximum pupil-size. This is because the PONE is concerned with increases in the maximum pupil-size in response to stimuli – by measuring the maximum size during the baseline, the likelihood that any baseline to trial difference is simply the difference between the pupil at minimum and maximum amplitude during hippus is reduced.

Gaze position affects the size of the pupil as perceived by eye-trackers such as the EyeLink, which measure pupil-size in eye-tracker camera pixels (Pomplun, & Sunkara, 2003; Pomplun, Sunkara, Fairley, & Xiao, 2009) (Tobii eye-trackers

measure the length of an ellipse fitted to the pupil which is less distorted by perspective; Klingner, 2010). Due to effects of gaze position on the EyeLink, in this thesis all events were presented in the centre of the monitor. Each trial did not begin until the participant had fixated the centrally-positioned drift correction dot, which was then followed by a fixation cross. To prevent luminance changes, which could trigger the PLR (see section 1.2.1.1), an isoluminant mask consisting of “&&&&&&” or “HHHHHH” (matched with stimuli for character length) preceded each stimulus. This was followed by the stimulus, which either remained in position for the duration of the trial, or was replaced by the isoluminant mask. Within an experimental condition stimuli were the same number of characters, presented in a Monospaced font, and all words subtended no more than 3° of visual angle to ensure they fell within the fovea, reducing the likelihood that participants would need to make a second fixation to read the stimulus and induce local luminance changes or distortions in pupil shape. Stimuli were achromatic, stationary and of constant contrast in order to control for pupil-size changes in response to visual stimulus features, and participants were asked to remain still during the experiment to prevent accommodation-related changes (Loewy, 1990).

Another source of noise are blinks and the lid-closure reflex (see Chapter 1 section 1.1.2.3), which causes both pupils to briefly contract and redilate. In a methodology paper, Nakayama (2006) found that blinks had a significant effect on both pupil-size and Pupil Unrest Index (PUI) when participants carried out a mental arithmetic task, and that estimation of pupil-size during blinks provided a pupil grand average that was more sensitive to the experimental manipulation in a small sample size ($n=5$). An alternative method to blink estimation or correction (Klingner, 2010), and the one used in this thesis, is blink suppression – asking participants to try to blink only between trials, as trials were only 2000ms long. The experimenter could see the eye

image during the experiment and wait until after a blink occurred to trigger the next trial. Blink reduction was especially important in Experiment 8, which used ERP measures, as eye blinks and eye movements have a detrimental effect on EEG recording (see section 2.2.1) due to large electrical signals produced by the eye muscles, and whilst random blinks will average out, stimulus-linked blinks will average into the grand average.

By pooling data from 20,000 binocular blinks Klingner (2010) asserted that stimulus-linked blinks were associated with a reduction in pupil-size in the subsequent 1000ms by $\sim 0.03\text{mm}$, and an increase in pupil-size between 1000-2000ms by $\sim 0.05\text{mm}$ (Klingner, 2010). However, increased cognitive load is known to be associated with both higher blink rates and increased pupil-size, so this is not surprising (see Chapter 1, section 1.2.4.2). As stimuli were presented on all trials, the procedure was the same for old and new items, and blinks were minimised as far as possible, it is unlikely that blinks accounts for the difference in pupil-size for old and new items. To check this, the number of blinks made during each trial was automatically recorded and output alongside the pupil-size data. Paired-sample t-tests on blink rate between conditions were performed across all experiments and no significant differences in blink rates for old and new items were found.

2.1.3. Pupil-Size Analysis

During an experiment the EyeLink records raw data every 1-2ms (depending on sampling rate) including a timestamp, the X and Y position of the eye(s) being tracked in screen pixel co-ordinates, pupil-size and event-related messages signalling when the display software has reached particular points in the experiment, for example mask and stimulus onset and offset. Raw data is imported into Data Viewer (SR Research, Ontario, Canada), which allows the specification of time windows for the

extraction of calculated variables, such as maximum pupil-size, into summary trial reports which then were analysed in Excel (Microsoft) and SPSS 18 (IBM).

As seen in Chapter 1, task-evoked increases in pupil-size are usually less than 0.5-1.0mm (Klingner, 2010) or 10-20% of baseline (Beatty, 1982), which is equivalent in magnitude to the constant background variation caused by other influences such as hippus. This makes it virtually impossible to identify the task-related signal from noise on any individual trial. One method of enhancing the signal-to-noise ratio is to average multiple repeated trials of the same task (Beatty, & Lucero-Wagoner, 2000; Pomplun, & Sunkara, 2003), leading to consistent task-evoked responses averaging in, whilst random variations such as hippus should average out. This is the approach taken in this thesis; the mean number of trials per participant per condition was 33.9 old (range = 20.5-58.5, $SD = 4.65$) and 37.7 new (range = 20.6-73.1, $SD = 4.29$).

Within experiments, statistical comparisons of pupil-size for old and new items were made for all trials and/or only correct trials for old and new items. The analyses restricted to the correct responses allowed us to be sure that any differences in pupil size between old and new items were not due to some “error” response that may occur when participants realise they have made an incorrect response. In some instances it was not appropriate to analyse only correct trials, for example in Experiment 6 where participants were randomly preloading answers, or in Experiment 5 where they were instructed to say “new” to all items. Unfortunately it was not possible in most cases to analyse changes in pupil-size associated with incorrect responses, even though previous research has shown an interesting effect of an intermediate pupil-size for false alarms (Otero et al., 2011) – as insufficient false alarms and misses were made by participants to produce a meaningful average for analysis.

Other techniques that researchers have used to analyse pupil data include waveform analysis (e.g., Kuipers, & Thierry, 2011), wavelet transforms or decomposition to find brief discontinuities that differentiate cognition from light reflexes (e.g., Leal, Neves, & Vieira, 2011; Marshall, 2002; 2007), frequency-domain analysis (Kumar, n.d., cited in Klingner, 2010; Moloney et al., 2006; Nakayama, & Shimizu, 2004), principle and independent component analysis (Jainta, & Baccino, 2010), analysis of average pupil-size (e.g., Klingner, 2010), and analysis of area under the pupil-response curve (Webb, Honts, Kircher, Bernhardt, & Cook, 2009). Oliveira et al. (2009) used Principle Component Analysis (PCA) to isolate changes in pupil diameter due to the local luminance changes from changes due to stimuli in a web search task. Jainta and Baccino (2010) used PCA and Independent Component Analysis (ICA) to reveal the main and hidden contributions to pupil responses from participants who were reading or performing easy or difficult mental arithmetic. They identified three components in the individual pupil responses, only one of which changed with task difficulty and accounted for 50% of variance during the most difficult task. Jainta and Baccino (2010) proposed that this component might be mental effort, but did not speculate as to the nature of the other two components.

The focus of the present thesis was to characterise the recently identified PONE in terms of the cognitive processes that may underlie it. As such, the relatively straightforward PDR was used as the methods described above are more suited to characterising the **nature** of the pupil response itself, possibly with a view to exploring its neural underpinnings.

2.2. Event-Related Potentials (ERPs)

2.2.1. ERP Data Acquisition

Event-Related Potentials (ERPs) are averaged waveforms identified as positive or negative deflections of the Electroencephalograph (EEG) voltage (Luck, 2005; see Chapter 6, section 6.1.1 for further discussion). Measured from the scalp, EEG recordings are made using arrays of electrodes in predetermined positions, covering the majority of the participants head, often as part of a net or cap. The traditional international 10–20 and 10–10 electrode configurations have 22 and 42 electrodes respectively (see Figure 2-5; Jasper, 1958; Michel et al., 2004; Pivik et al., 1993).

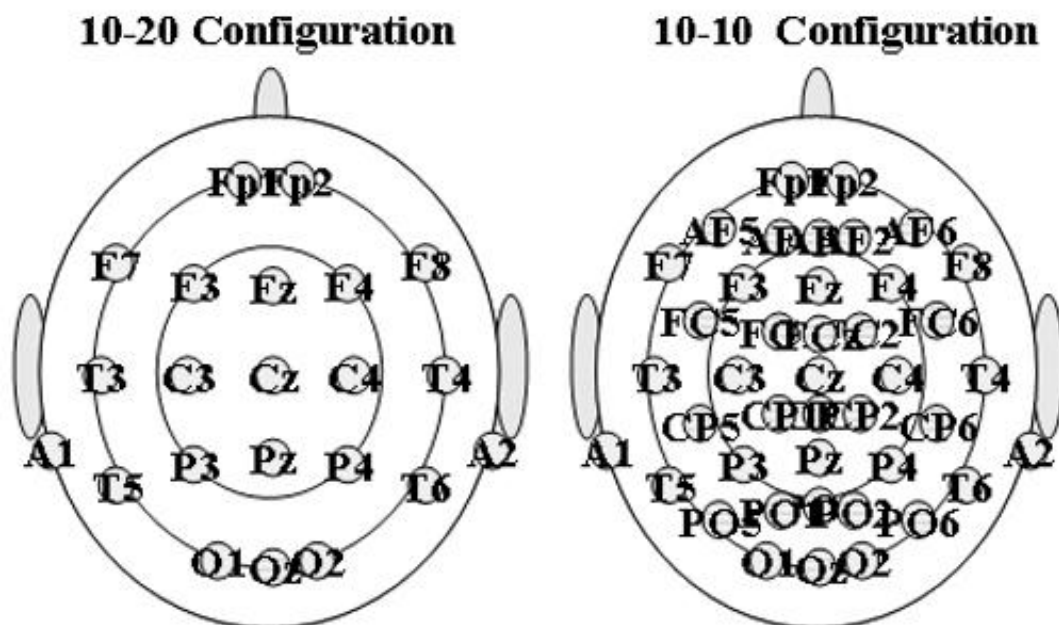


Figure 2-5: Traditional 10-20 and 10-10 electrode configurations (adapted from Reynolds, & Richards, 2009).

“Geodesic sensor nets”, such as the Electrical Geodesics Inc. (EGI) nets used in Experiment 8 of this thesis (see Chapter 6, section 6.2), have a high-density (or dense-array) electrode configuration of 64, 128, or 256 equidistant electrodes, approximately 35-40 mm apart (for adults, depending on head size and net size)

covering most of the scalp surface (Electrical Geodesics Inc; Tucker, 1993; Tucker, Liotti, Potts, Russell, & Posner, 1994, see Figure 2-6 and Figure 2-7).

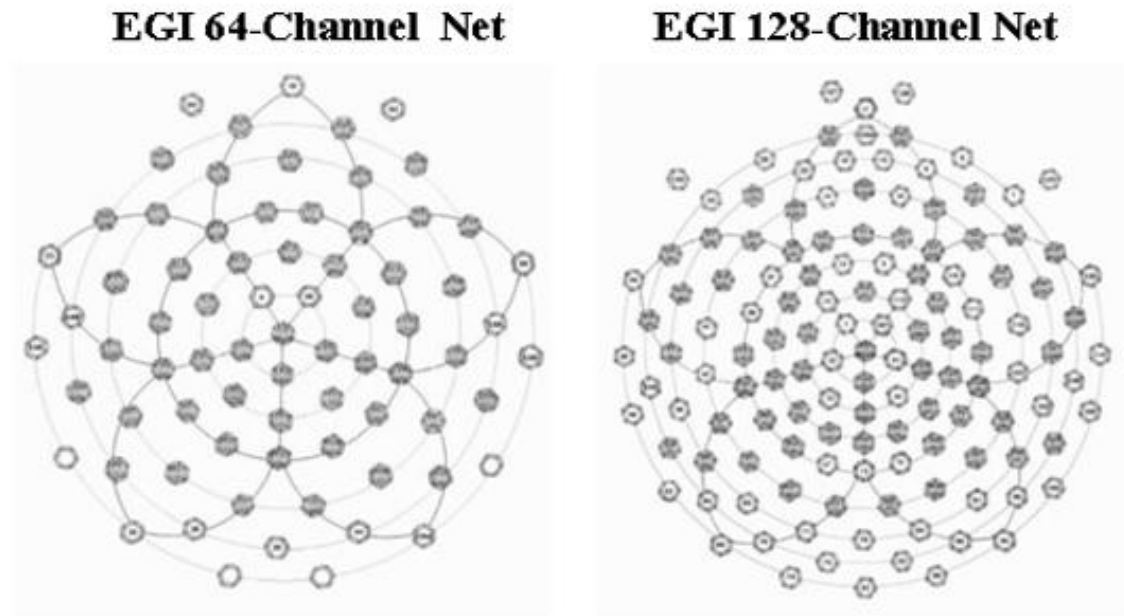


Figure 2-6: Geodesic sensor net 64 and 128 channel electrode maps (adapted from Reynolds, & Richards, 2009).



Figure 2-7: 128 channel Geodesic sensor net worn by models.

The traditional 10-20 positions have been updated to replace T3/T4 with T7/T8, and T5/T6 with P7/P8 in line with guidelines issued by the American Electroencephalographic Society (1991; 1994), and, to make room for P9/P10, electrodes P7/P8 were moved to a more superior site (see Figure 2-8).

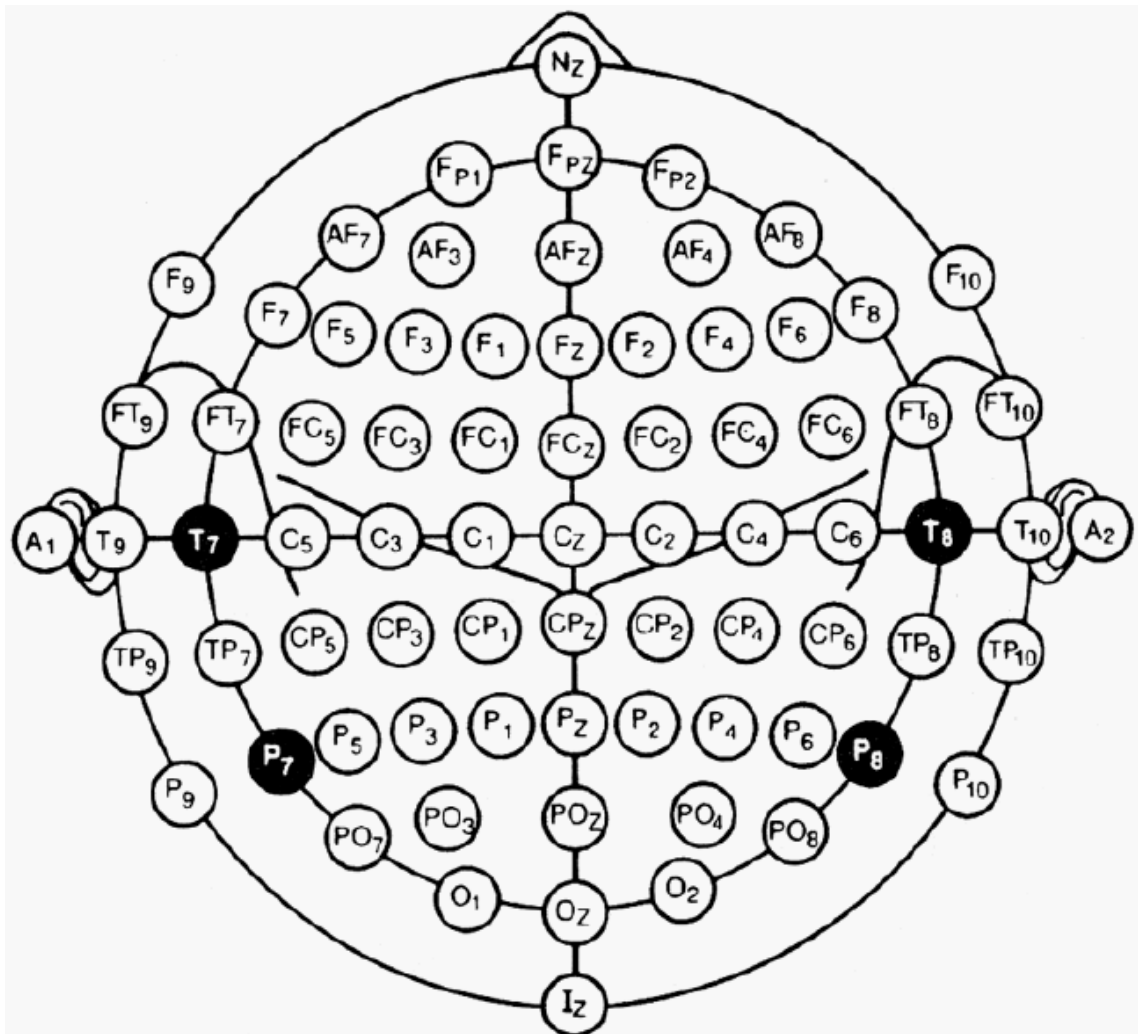


Figure 2-8: Modified combinational nomenclature for the 10-10 system (from the American Clinical Neurophysiology Society, 2006).

The geodesic configuration differs from the electrode placement sites of the International 10-10 and 10-20 systems, but an approximate correspondence between the two has been established (Luu, & Ferree, 2000; Srinivasan, Tucker, & Murias, 1998). Luu and Ferree (2000) computed corresponding positions between the two systems using maximum arc length distance of 0.20 of the radius (see Figure 2-9).

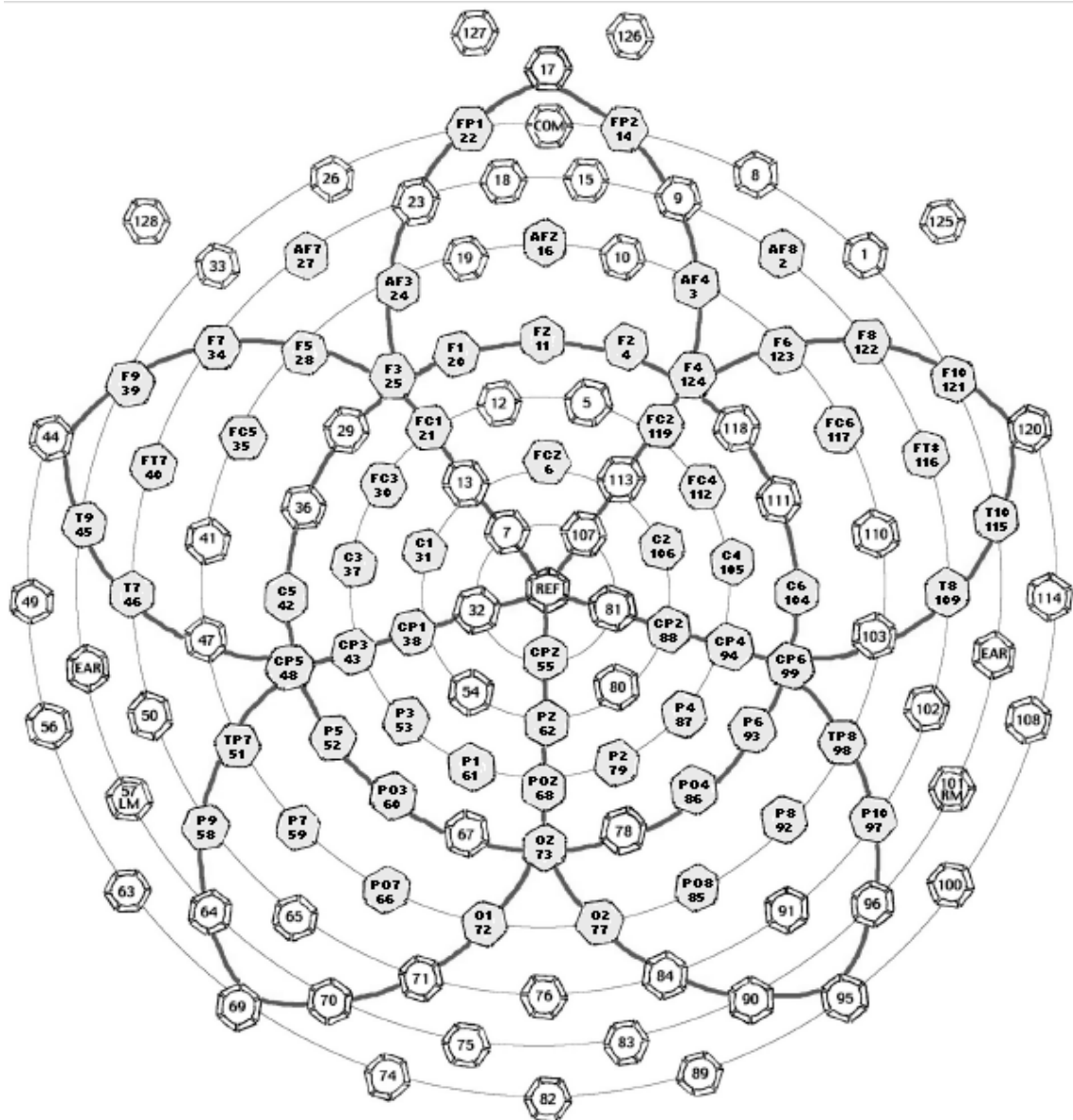


Figure 2-9: The 10-10 system overlaid on the Geodesic 128 sensor net (adapted from Luu, & Ferree, 2000).

The changes in voltage measured by EEG equipment are tiny (often less than $10\mu\text{V}$), and typically smaller than the background electrical fluctuations from electronic equipment (including experimental apparatus), skin potentials, muscle and eye-movements (Luck, 2005). In order to accurately measure dipoles at the surface of the scalp the signal must be amplified 10,000-50,000 times, along with the noise, therefore it is important to reduce sources of noise as far as possible. Ways of doing this include shielding electronic equipment whilst keeping it as far from the participant

as possible, reducing the impedance of the skin via abrasion or using a high-impedance system, asking participants to relax so that they are not clenching their facial or neck muscles, and asking participants to fixate the centre of the screen and only blink between trials (Luck, 2005). As occurred in Experiment 8, participants can be seated in a Faraday cage, which shields the entire experimental setup from outside electromagnetic radiation; however care must be taken with any equipment (e.g., monitors, eye-trackers) used inside the cage as emissions will be trapped within the cage. Endogenous noise can also occur in the form of alpha-waves in participants who are tired, bored or sleepy, similar to hippus. This can be reduced by using well rested participants and offering rest breaks and water between blocks.

2.2.2. ERP Data Analysis

Analysis of EEG data was carried out by segmenting the continuous epoch into sections that began 200ms prior to stimulus presentation (-200ms) and ended 1000ms after stimulus presentation, using event markers communicated by the experimental software (E-Prime 2.0) to the EEG software (Net Station), and grouped according to condition. Grand-average waveforms were generated for each participant, baseline-corrected using the period -200 to 0ms, and re-referenced offline to average mastoid electrodes (Nunez, 1981) after these channels were verified as having made a good, relatively artefact-free recording (it is not possible to make this check for the online reference during recording). Average mastoid reference was selected as this is a commonly used reference, allowing comparison with other studies, it is also a convenient site that does not cause discomfort or distraction, provides good electrical conduction, and given that all references have their limitations it is as good a reference as any (Luck, 2005; however see also Dien, 1998).

Net Station Waveform Tools (Electrical Geodesics, Inc) were used to extract the mean amplitude for the windows of interest, and data were analysed using SPSS 18 (IBM). Data were analysed in their raw form, without normalisation conversion to relative differences as only one electrode factor had more than two levels, and although theoretically appealing, Urbach and Kutas (2002) state that normalisation fails to achieve the desired effect of removing significant condition by electrode site interactions (Luck, 2005). To reduce potential violations of sphericity, and retain topographical detail, lateral electrode position was analysed using two factors with two levels (hemisphere: left, right; site: superior, inferior), rather than one factor with four levels (Luck, 2005). Analyses included parallel strings of electrodes (e.g., MacKenzie, & Donaldson, 2007) rather than groups of electrodes (e.g., Curran, 2000).

In line with other ERP memory research (e.g., Curran, 2000; Finnigan et al., 2002; MacKenzie, & Donaldson, 2007), multiple univariate analyses were performed, rather than a single multivariate analysis (where Mauchly's test indicated that the assumption of sphericity had been violated, degrees of freedom were corrected using *Greenhouse-Geisser estimates of sphericity ($\epsilon < 0.75$) or *Huynh-Feldt estimates of sphericity ($\epsilon > 0.75$)). This was because whilst MANOVA would determine whether or not there was an effect, it would not reveal where, and would still require multiple follow up ANOVA. Kiebel and Friston (2004) have stated that multivariate and mass univariate are not dissimilar, and Groppe, Urbach and Kutas (2011) review four promising methods of mass univariate analysis that control for familywise error and are particularly suited for exploring ERP data.

Luck (2010) questions whether it is legitimate to compare grand averages of conditions containing different numbers of trials as is common in ERP experiments.

Issues may arise because waveforms formed from fewer trials will contain more noise due to a lower signal-to-noise ratio than averages of larger numbers of trials (Luck, 2010). Variation in noise is more of a concern when measuring peak amplitude rather than mean amplitude; it biases the measurement because a spurious peak has more influence over the final value of the peak measurement due to fewer contributing trials (Luck, 2010). The ERP experiment reported in this thesis is concerned with mean amplitude, which is an unbiased measure even when trial numbers differ, and so perfectly good trials do not need to be discarded simply to even the numbers (Luck, 2010).

2.3. Stimuli and Participants

2.3.1. Word Selection

With the exception of the artificial grammar condition of Experiments 3 and 4 (see Chapter 4), study and recognition lists for the experiments in this thesis were created using nouns selected from the MRC Psycholinguistic Database (Coltheart, 1981). Items within a list were matched for length (5, 6, or 7 letters long), and between lists were matched for lexico-semantic features such as frequency, familiarity and imageability, according to K-F norms (Kucera, & Francis, 1967), as these are known to affect both memory performance (e.g., Balota, & Neely, 1980; Bauer, Olheiser, Altarriba, & Landi, 2009; Deese, 1960; Gorman, 1961; Gregg, 1976; Schulman, 1967), and pupil-size (Colman, & Paivio, 1969; 1970; Kahneman, & Peavler, 1969; McElvain, 1970; Paivio, & Simpson, 1966; 1968; Simpson, & Paivio, 1968; see Chapter 1, section 1.2.3.6). For example, an item is more likely to be correctly recognised as “old” if it is relatively uncommon (Shepard, 1967), or if it is concrete rather than abstract (Gorman, 1961). Words with emotional or offensive content were excluded due to their potentially biasing effects on both memory (Bauer et al., 2009)

and pupil-size (Johnson, 1971; Stelmack, & Mandelzys, 1975; see Chapter 1, section 1.2.2). Different word lists were used for each experiment to prevent confounds if a participant took part in more than one experiment, which was particularly important for experiments with implicit tests of memory as performance on these has been shown to be more enduring than on explicit tests of memory (Allen and Reber, 1980).

2.3.2. Selection of Participants

Participants were required to be native English speakers, and to have normal, or corrected-to-normal, vision in at least one eye, with glasses or contact lenses to be brought to the experiment if required. Whilst both contact lenses and glasses have effects on the refraction of infra-red light (Dahlberg, 2010; Wang, 2010), experiments in this thesis involve participants fixating the centre of the screen, minimising artifacts normally associated with gaze-tracking. Participants were prevented from participating in to both Experiments 3 and 4, or in both Experiments 8 and 9, as the stimuli used were identical.

Although resting pupil-size decreases with age, pupil responses appear to remain relatively unchanged during adulthood (see Chapter 1, section 1.1.1; Kim, Beversdorf, & Heilman, 2000; Kumnick, 1956; Porter et al., 2010). However, the correlation between pupil-size and age is slightly lower in psychiatric populations than healthy controls, whereby resting pupil-size in participants with mental health problems does not decrease with age as much as for healthy participants, possibly due to comorbid anxiety (Liakos, & Crisp, 1971). Even treated and remitted schizophrenics have abnormal pupil responses, such as decreased dilations to stimuli and faster working memory overload compared to controls (Andreassi, 2000; Granholm, & Verney, 2004; Minassian, Granholm, Verney, & Perry, 2004). Therefore it was relatively important to keep age constant as baseline measurements are used to calculate PDR, and it was

important for the present experiments to recruit participants without significant mental health difficulties. The average age of 377 participants (113 male) across all experiments was 24.4 years ($SD = 6.9$ years).

As pupil-size is influenced by thoughts and feelings, including physical sensations such as pain or discomfort, and background noises and distractions, care was taken to seat participants comfortably in an adjustable chair, maintain the laboratory at an adequate temperature, provide water and breaks if required and remove sources of distraction.

3. Replicating and Extending the Pupil Old/New Effect

In Chapter 1, the pupillometry literature and relevant recognition memory literature was reviewed, showing that since the 1960s a number of studies have, directly or indirectly, measured the effect of encountering novel versus learned stimuli on pupil-size (e.g. Bradshaw, 1967; Bradshaw, 1968; Garrett, Harrison, & Kelly, 1989; Gardner et al., 1974a; 1974b; Maw, & Pomplun, 2004; Otero et al., 2006; 2011; Vö et al., 2008). It is clear that the Pupil Old/New Effect (PONE) is now well established for “explicit” recognition, but what is not clear is exactly what the effect represents, whether it is associated with specific mnemonic processes, or to what extent the PONE is linked to conscious awareness. As yet no research has investigated whether the PONE can also be observed when memory is tested “implicitly”. The aims of the two experiments presented here were to replicate the PONE in a standard explicit memory recognition test, and explore whether a similar effect can be observed in an “implicit” recognition test.

A large body of literature suggests that a distinction can be made between explicit and implicit memory. Implicit memory is defined experimentally as a change in performance that results from previous exposure to items, but in the absence of a conscious recollective experience of the exposure itself (Dienes, & Berry, 1997; Stevens, Wig, & Schacter, 2008). One line of research that has been used to support the distinction between implicit and explicit memory is experimental dissociations in healthy participants – manipulations which affect performance on one but not the other type of memory task (see Foster, & Jelicic, 1999). For example, a Levels Of Processing (LOP) manipulation (Craik, & Lockhart, 1972) enhances performance on explicit tests of memory for items processed more deeply during study (semantic processing; e.g., “Generate a sentence using this word”) compared

to items processed shallowly during study (surface processing; e.g., “What colour is the text?”), but does not differentially affect performance when memory is tested implicitly, with techniques such as perceptual identification (Jacoby, & Dallas, 1981), word stem completion (Graf, Mandler, & Haden, 1982), and word fragment completion (Roediger, Weldon, Stadler, & Riegler, 1992). In contrast, “surface” manipulations (such as keeping the font of items constant between study and test) enhance performance on implicit tests of memory, but do not affect explicit tests of memory (Stevens, Wig, & Schacter, 2008).

In addition, a growing literature demonstrates implicit-explicit dissociations in neuropsychological patients, including relatively intact implicit memory in patients with amnesia (e.g., Laeng et al., 2007; Verfaellie, Bauer, & Bowers, 1991; for a review see Schacter, McAndrews, & Moscovitch, 1988). For example, Nissen and Bullemer (1987; see also Nissen, Willingham, & Hartman, 1989) showed that when presented with a ten-trial repeating light sequence, which the participants then had to recreate, the performance of participants with Korsakoff’s amnesia improved as the sequence was repeated, consistent with performance of control participants, even though unlike controls the Korsakoff’s participants were not consciously aware of the pattern.

As discussed in Chapter 1, section 1.3.1.2, familiarity as measured by “know” responses (Gardiner, 1988; Tulving, 1985a) has been shown to respond in a similar manner to so called “implicit” memory in a variety of experimental manipulations (Paller, Voss, & Boehm, 2007; Yonelinas, 2002). For example, priming manipulations where the stimulus is briefly presented prior to testing lead to feelings of familiarity (Jacoby, & Whitehouse, 1989), which enhanced “know” responses, without influencing “remember” responses (Huber, Clark, Curran, & Winkielman, 2008). Similarly, familiarity and recollection often dissociate in neuropsychological patients

with amnesia, Parkinson's disease or Alzheimer's (Cohn et al., 2010; O'Connor, & Ally, 2010; Weiermann, Stephan, Kaelin-Lang, & Meier, 2010; Yonelinas et al., 1998).

To date, only one study appears to have assessed the relationship between pupil-size and recognition memory performance in amnesic patients. Laeng et al. (2007) tested two patients with amnesia by reading unfamiliar or fictional short facts (e.g., "penguins lay blue eggs"), while an image related to a word in the sentence (e.g., "eggs") was presented in one of four boxes on a computer screen. The patients were asked questions based on the facts (e.g., "what colour are penguins' eggs?"), and despite answering very few questions correctly, their eyes focused on the box in which the relevant image had been presented. The authors argued that this finding suggests the patients had an implicit memory for the location of the picture. In a second experiment, Laeng et al. (2007) carried out a picture based old/new recognition task with three amnesic patients. One patient answered "new" to every question, whilst the others made 46.6% and 70.2% correct decisions. Interestingly, in contrast to most recent research which has found significantly larger pupil-sizes for old words, the amnesic patients' pupil-sizes were greater for new than old words. The reasons for this discrepancy are not clear, but given their amnesic status, it might be argued that any correct recognition would be based on implicit or non-recollective memory processes (suggesting that the standard PONE reflects primarily recollective mnemonic processes).

There are a number of issues which hinder interpretation of the Laeng et al. (2007) study. The sample was very small, the amnesic patients had different aetiologies and, as is clear from their performance, a wide range of memory difficulties. Importantly, baseline pupil measurements were only taken from a single participant during the blank screen between pictures. Stimuli were a mixture of colour pictures

and photographs of objects and faces, and only one participant had not seen the same set of stimuli in the two preceding experiments. Areas of the brain have been shown to respond to stimulus novelty (Habib, & LePage, 1999; Tulving, & Kroll, 1995), showing decreased activation to repeated presentation (repetition suppression; e.g., Schacter, & Buckner, 1998; Buckner et al., 1998; Grill-Spector, & Malach, 2001; van Turennout, Ellmore, & Martin, 2000), and some studies have demonstrated a “novelty” pupil response (Andreassi, 2000; Janisse, 1977), which may be part of an orienting response to salient stimuli (Lynn, 1966; Pavlov, 1927; Sokolov, 1963). It is therefore possible that in the absence of a PONE in amnesic participants, the novelty response is instead the most visible influence on pupil-size between new and old items.

Implicit tests of memory offer a way to study the influence of familiarity on recognition decisions in the absence of recollective processes, and LOP manipulations (e.g., Craik, & Lockhart, 1972) have been shown to have different effects on explicit and implicit tests of memory (e.g., Jacoby, & Dallas, 1981). The first experiment combined a LOP manipulation at study with explicit and implicit tests of recognition memory. Its aim was to replicate the PONE, and determine whether the PONE can also be observed when participants are exposed to novel and learned items, but are not asked to make a recognition decision based on conscious recollection. To this end, a standard recognition task was used in one condition (called the “explicit” condition), and a perceptual fluency recognition task (as used by Jacoby and Dallas, 1981) was used in the second condition (hereafter referred to as the “implicit” condition). In perceptual fluency tasks, the recognition of very briefly presented items is facilitated if they have previously been encountered during the learning phase, without participants necessarily being able to consciously recollect the initial learning experience. The LOP manipulation employed at learning was included in an attempt

to replicate Otero et al.'s (2011) finding that the PONE was larger for items which had been encoded with deep orienting instructions compared to those encoded with shallow orienting instructions, and larger for shallow items than new items.

On the basis of previous research, pupil-size was expected to increase for old compared to new items in the explicit recognition test. However, Võ et al.'s (2008) cognitive load account of the PONE (see Chapter 1, section 1.2.2.1) does not consider whether the PONE should also be observed when participants recognise previously encountered stimuli based on a familiarity judgement rather than recollection. If a PONE was observed during a recognition judgment made on the basis of familiarity alone, this would undermine the claim that pupil dilation reflects a) recollective processes and b) cognitive load. Due to their emphasis on conscious recollection, Võ et al.'s (2008) cognitive account would predict that the PONE in the explicit condition should be smaller for deeply encoded items (less effort needed for recollection) than shallowly encoded items, whereas Otero et al.'s (2011) "memory signal" explanation predicts that deeply encoded items should elicit a larger pupil-size (stronger memory signal) than shallowly encoded items, which should elicit a larger pupil-size than new items.

It was predicted that the PONE would not be observed for either semantic or shallow items in the implicit recognition task because conscious recollective processes would not be involved in perceptual priming. If, however, the absence of PONE allowed the novelty pupil effect to dominate pupil-size (as may have occurred in Laeng et al., 2007), it was predicted that new items might elicit a larger pupil-size than surface and semantic items in the implicit condition. An open prediction was made as to whether there would be a difference between surface and semantic items in the implicit condition. An LOP effect was also predicted between conditions in the behavioural

data whereby a higher number of semantic (deeply encoded) items would be correctly recognised in the explicit condition due to deep encoding, and a higher number of surface (shallowly encoded) items would be correctly recognised in the implicit condition due to perceptual fluency.

3.1. Experiment 1 – Implicit vs. Explicit Tests of Recognition

3.1.1. Method

3.1.1.1. *Participants*

Fifty participants (20 male; age range: 18.4-36.5, $M = 23.4$, $SD = 4.01$) with normal or corrected-to-normal vision in at least one eye, were recruited from the psychology course-credit and subject pools at the University of Sussex, and through personal contact. Participants were briefed with a detailed consent form (specific to the condition to which they were allocated) and verbal description of the methods and procedure, and were invited to ask questions. Participation was on a voluntary basis, and participants were thanked and debriefed at the end with a verbal description of the aims of the study and the opportunity to ask further questions. The experiment was approved by the relevant ethics committee.

3.1.1.2. *Materials/Apparatus*

Two word lists were created using nouns selected from the MRC Psycholinguistic Database. The learning list comprised 40 items, whilst the recognition list contained those 40 nouns plus 40 new items. All items were 6 letters long, and lists were matched for concreteness (learning items range = 305-634, $M = 523$, new items range = 296-635, $M = 525$), familiarity (learning items range = 436-621, $M = 547$, new

items range = 428-632, $M = 549$), imageability (learning items range = 368-643, $M = 551$, new items range = 324-646, $M = 552$) and frequency (learning items range = 21-348, $M = 95$, new items range = 18-492, $M = 95$), according to the K-F norms. Words likely to elicit a strong emotional response were removed. The learning list and recognition test were presented in black 20pt Arial font in the centre of a light grey background under fixed illumination. Items were presented using the Experiment Builder software associated with the EyeLink II eye-tracker (SR-Research, Ontario). All items are presented in Appendix A. During the recognition test, pupil-size was recorded using an EyeLink II head-mounted eye-tracker with a temporal resolution of 2ms and a spatial resolution of around 0.25 degrees. The stimuli were displayed on a 21 inch CRT monitor with a screen resolution of 1,280 x 1,024 pixels and a refresh rate of 60Hz. Actual screen dimensions were 40cm horizontal and 30cm vertical. Participants were seated approximately 70cm from the screen in an adjustable chair that had been modified to prevent any rotational movement.

3.1.1.3. *Design and Procedure*

The experiment comprised two conditions, an explicit recognition condition and an implicit recognition condition, and in a between-subject design half of the participants completed each condition. Both conditions contained a learning and recognition phase. The learning phase was identical between the conditions. The 40 learning list items were presented on screen for 3000ms. Before each item was presented, participants saw a screen instructing them to process the following item at either a surface level ("How many vowels in...") or semantic level ("Give me a synonym for..."). The same items were associated with the same LOP (shallow or deep) for all participants across both conditions, and an equal number of items were processed at the deep and shallow level. Participants were required to give an answer for each question.

For the explicit recognition condition, the 80 recognition list items were presented on screen for 2000ms after which the participant was prompted to say whether the item was old (previously encountered in the learning phase) or new (not previously encountered). The next screen required participants to estimate confidence in their decision with a number between 1 and 5, where 1 represented a complete guess and 5 represented total confidence. This screen was then replaced by a drift-correction dot in the centre of the screen before presentation of the next item. Old/new judgements and confidence estimates were recorded on the computer after each recognition item.

For the implicit recognition condition, the 80 recognition list items were present for two monitor-refresh cycles (33.3ms at 60Hz), as determined by the eye-tracker software (Experiment Builder, SR Research). In order that participants were looking at the item during its brief presentation, participants were asked to blink whilst the drift-correct dot was on screen and state when they were ready to proceed without blinking for a few seconds. They were then required to read the item aloud and their answers were noted by hand and entered into the computer at a later stage.

3.1.1.4. Pupil Recording

Maximum pupil-size was recorded from the right eye during each recognition period – the time during which the item was on screen for the explicit condition, and the time which the item and the re-mask were on screen for the implicit condition. A Pupil Dilation Ratio (PDR; see Chapter 2, section 2.1.2.1) was calculated expressing the maximum pupil-size for each 2000ms recognition trial as a proportion of the maximum pupil-size during that trial's 200ms baseline.

3.1.2. Results

3.1.2.1. Behavioural Data: Old/New Responses

The proportion of correct responses to old and new items was calculated for implicit and explicit conditions. A 2 x 2 ANOVA with within-subject factor of item-type (old vs. new) and between-subject factor of condition (implicit vs. explicit) showed a significant main effect of condition ($F(1,48) = 21.1$, $MSE = 0.013$, $p < .001$, $\eta_p^2 = .305$) – more correct responses were made in the implicit condition than the explicit condition. This main effect was qualified by a significant item-type by condition interaction ($F(1,48) = 27.9$, $MSE = 0.012$, $p < .001$, $\eta_p^2 = .368$) – participants responded correctly more often to old items ($M = .947$, $SD = 0.074$) than new items in the implicit condition ($M = .840$, $SD = 0.142$, $t(24) = 5.36$, $p < .001$, $r = .545$), but responded correctly more often to new items ($M = .851$, $SD = 0.115$) than old items in the explicit condition ($M = .723$, $SD = 0.112$, $t(24) = -3.22$, $p < .01$, $r = .302$). The main effect of item-type was not significant ($F(1,48) = 0.220$, $MSE = 0.012$, $p > .05$, $\eta_p^2 = .005$; see Figure 3-1).

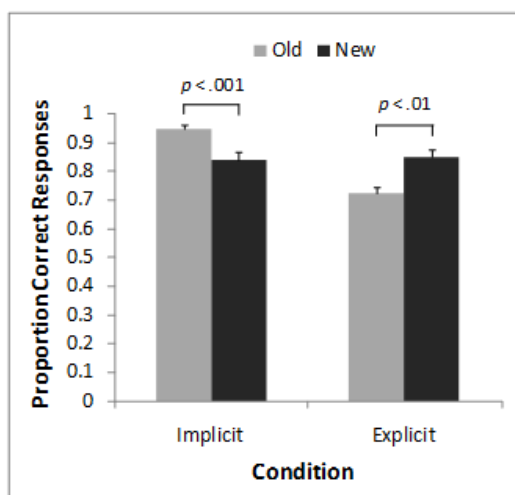


Figure 3-1: Proportion of correct responses in each condition. Error bars show standard error of mean.

In order to determine the effect of the LOP manipulation on recognition memory, the proportion of correct responses to surface and semantic old items were analysed in a 2 x 2 mixed ANOVA with LOP (surface vs. semantic) as a within-subject factor and condition (implicit vs. explicit) as a between-subject factor. The main effect of LOP was significant ($F(1,48) = 135.7$, $MSE = 0.012$, $p < .001$, $\eta_p^2 = .739$) – more correct responses were given to semantic items than to surface items. This main effect was qualified by a significant LOP by condition interaction ($F(1,48) = 113.6$, $MSE = 0.012$, $p < .001$, $\eta_p^2 = .703$) – participants responded correctly much more often to semantic items ($M = .970$, $SD = 0.035$) than surface items in the explicit condition ($M = .476$, $SD = 0.217$, $t(24) = 11.6$, $p < .001$, $r = .848$) whereas in the implicit condition the proportion of correct responses to semantic items ($M = .958$, $SD = 0.064$) and surface items ($M = .936$, $SD = 0.093$) were more similar and only approached significance ($t(24) = 1.90$, $p = .07$, $r = .131$). The main effect of condition was also significant ($F(1,48) = 69.0$, $MSE = 0.018$, $p < .001$, $\eta_p^2 = .590$; see Figure 3-2).

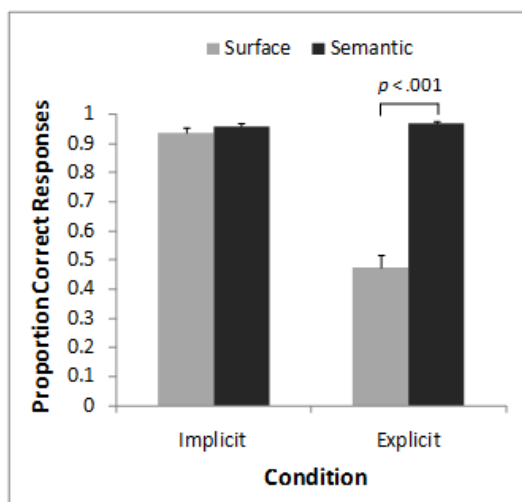


Figure 3-2: Proportion of correct responses to surface and semantic items in each condition. Error bars show standard error of mean.

3.1.2.2. Behavioural Data: Confidence

To determine the relationship between confidence and performance, the average participant-reported confidence was calculated for the explicit condition (confidence estimates for the implicit condition was not measured because participants were not making a recognition judgement). Confidence ratings were analysed in a repeated measures ANOVA with within-subject factors of item-type (old vs. new), and response (old vs. new), which showed a significant main effect of item-type ($F(1,17) = 98.1$, $MSE = 0.094$, $p < .001$, $\eta_p^2 = .852$) – average confidence for old items was higher than for new items. This was qualified by a significant item-type by response interaction ($F(1,17) = 118.4$, $MSE = 0.156$, $p < .001$, $\eta_p^2 = .874$) – average confidence was higher for old items given an old response ($M = 4.44$, $SD = 0.324$) than a new response, and average confidence was higher for new items given a new response ($M = 3.71$, $SD = 0.794$) than an old response. The main effect of response was not significant ($F(1,17) = 0.001$, $MSE = 0.441$, $p > .05$, $\eta_p^2 < .001$; see Figure 3-3).

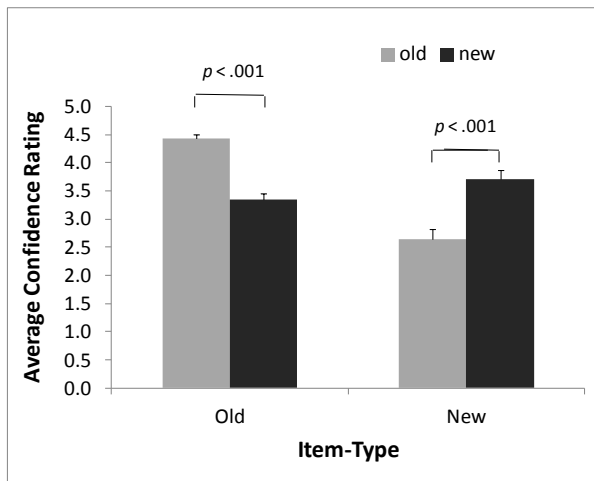


Figure 3-3: Average confidence rating for old and new responses to old and new items in the explicit condition. Error bars show standard error of mean.

3.1.2.3. Pupil-Size Data

Average PDR for old and new items was calculated for the implicit and explicit conditions. As PDR is a function of baseline pupil-size, baseline pupil-sizes to old and new items in both conditions were compared to ensure that any differences in PDR were not due to baseline differences. The difference was not significant ($F(1,48) = 0.908$, $p > .05$, ns, $\eta_p^2 = .019$). A 2 x 2 ANOVA of PDR with within-subject factor of item-type (old vs. new) and between-subject factor of condition (implicit vs. explicit) showed that the main effect of item-type was not significant ($F(1,48) = 0.63$, $MSE < 0.001$, $p > .05$, $\eta_p^2 = .013$), neither was the main effect of condition ($F(1,48) = 0.74$, $MSE = 0.011$, $p > .05$, $\eta_p^2 = .015$), however the interaction between item-type and condition was significant ($F(1,48) = 18.9$, $MSE < 0.001$, $p < .001$, $\eta_p^2 = .282$). As predicted, average PDR was larger for old items ($M = 1.160$, $SD = 0.069$) than new items ($M = 1.150$, $SD = 0.065$, $t(24) = 2.71$, $p < .01$, $r = .234$) in the explicit condition, and was larger for new items ($M = 1.180$, $SD = 0.081$) than old items ($M = 1.165$, $SD = 0.080$, $t(24) = -3.40$, $p < .01$, $r = .325$) in the implicit condition (see Figure 3-4).

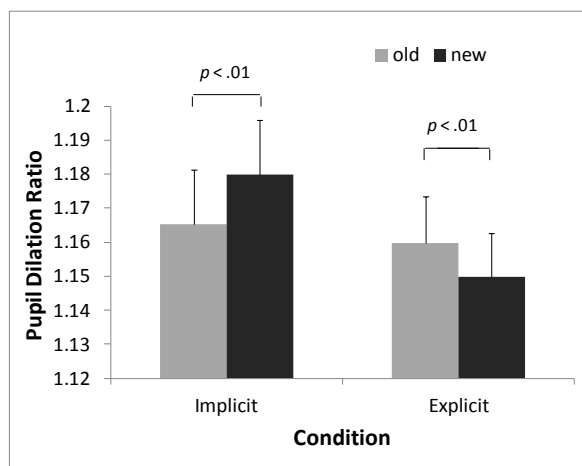


Figure 3-4: Pupil dilation ratio for old and new items in each condition. Error bars show standard error of mean.

The ANOVA was repeated, but with the data averaged across only those trials to which participants gave correct responses. An identical pattern of results was found (perhaps unsurprisingly given the high level of accuracy with which both tasks were completed). The main effect of item-type was not significant ($F(1,48) = 2.49$, $MSE < 0.001$, $p > .05$, $\eta_p^2 = .049$), neither was the main effect of condition ($F(1,48) = 0.37$, $MSE = 0.011$, $p > .05$, $\eta_p^2 = .008$), however there was a significant interaction between item-type and condition ($F(1,48) = 34.6$, $MSE < 0.001$, $p < .001$, $\eta_p^2 = .419$).

To determine whether pupil-size was influenced by the LOP manipulation, a 2 x 3 ANOVA on mean PDR values for correct items, with within-subject factor of LOP (new vs. surface vs. semantic) and between-subject factor of condition (implicit vs. explicit) was performed. There was a significant main effect of LOP ($F(1.75,82.3) = 8.99$, $MSE = 0.001$, $p < .001$, $\eta_p^2 = .161$; Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = 7.05$, $p < .05$), therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\epsilon = 0.97$)) – average PDR for semantic items was larger than for surface items or new items. This main effect was qualified by a significant LOP by condition interaction ($F(1.75,82.3) = 7.64$, $MSE < 0.001$, $p < .001$, $\eta_p^2 = .140$; Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = 5.97$, $p < .05$), therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\epsilon = 0.98$)) – average PDR for new items was smaller than for surface and semantic items in the explicit condition, whereas in the implicit condition, PDR for new items was larger than for surface or semantic items. The main effect of condition was not significant ($F(1,47) = 0.285$, $MSE = 0.016$, $p > .05$, $\eta_p^2 = .006$; see Figure 3-5).

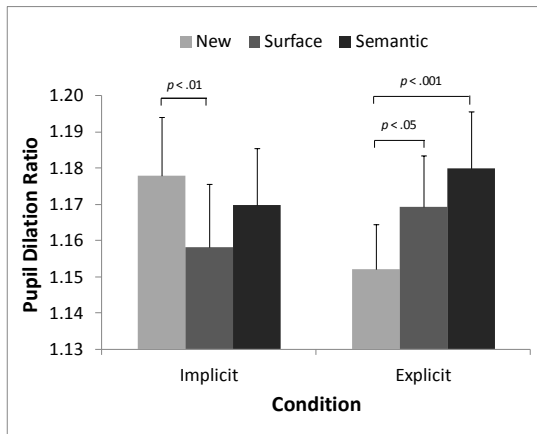


Figure 3-5: Pupil dilation ratio for new, surface and semantic items in each condition. Error bars show standard error of mean.

A priori t-tests revealed that, as predicted, in the explicit condition average PDR for correct semantic items ($M = 1.173$, $SD = 0.078$) was significantly larger than for correct new items ($M = 1.146$, $SD = 0.062$; $t(24) = 4.28$, $p < .001$, $r = .433$), as was average PDR for correct surface items ($M = 1.169$, $SD = 0.068$; $t(24) = 2.33$, $p < .05$, $r = .198$), and average PDR for correct semantic items was larger than for correct surface items at trend levels ($t(24) = 1.95$, $p = .08$). In the implicit condition, average PDR for correct new items ($M = 1.178$, $SD = 0.081$) was larger than for correct surface items ($M = 1.158$, $SD = 0.088$; $t(24) = 3.00$, $p < .01$, $r = .272$), however the differences between correct new and correct semantic, and correct surface and correct semantic items were not significantly different ($ts < 1.6$, ns).

3.1.2.4. Pupil-Size Data: Confidence Analysis

Participants made higher confidence ratings on average to their correct “old” judgments compared to their correct “new” judgements in the explicit condition. If the increase in pupil dilation observed when participants view old compared to new items reflects some kind of “arousal” associated with being correct, we might expect a relationship between the average PDR for correctly identified old items and confidence estimates for those correctly recognised stimuli. It is important therefore

to establish the extent to which the increase in PDR is associated with the increase in confidence that is associated with giving an old compared to new response.

A 2 x 2 repeated measures ANOVA of mean PDR values for correct items, with within-subject factors of item-type (old vs. new) and confidence (high vs. low) showed a significant main effect of confidence ($F(1,21) = 11.2$, $MSE = 0.004$, $p < .01$, $\eta_p^2 = .347$) – average PDR was larger for items ranked with high confidence (4 or 5) than for items with low confidence (1-3). This was qualified by a significant item-type by confidence interaction ($F(1,21) = 1.87$, $MSE = 0.002$, $p < .05$, $\eta_p^2 = .140$) – despite being overall slightly less confident in their correct rejections than their correct recognitions, the increase in average PDR with increasing confidence was greater for old items than new items. The main effect of item-type was not significant ($F(1,21) = 1.09$, $MSE = 0.001$, $p > .05$, $\eta_p^2 = .049$). Analysis was restricted to the 22 participants who had at least 5 high and low confidence correct old and new judgements.

3.1.3. Discussion

The present experiment replicated the basic PONE effect when memory was tested explicitly, but interestingly there was no PONE in the implicit condition. In addition, overall PDR was larger in the implicit condition compared to the explicit condition. Given the experimental design, it is not possible to say whether either of these effects is due to differences in task requirements (reading vs. recognition) or duration of stimulus presentation (33ms vs. 2000ms).

In an attempt to clarify the results of Experiment 1, a “control reading” condition was carried out, which acted as an additional comparison. This experiment was included in order to provide an estimate of the effect on pupil-size of simply reading items

presented onscreen for 2000ms during the recognition phase, without the requirement of making an old/new judgment.

3.2. Experiment 1b – Reading Condition

3.2.1. Method

3.2.1.1. *Participants*

Twenty-five participants (6 male; age range: 18.92-41.33, $M = 24.84$, $SD = 5.45$), with normal or corrected-to-normal vision were recruited from the psychology course-credit and subject pools at the University of Sussex, and through personal contact.

Participants were briefed with a detailed consent form and verbal description of the experiment, and invited to ask questions. Written consent was obtained prior to testing and participants were fully debriefed at the end. The experiment was approved by the relevant ethics committee.

3.2.1.2. *Materials/Apparatus*

As for Experiment 1.

3.2.1.3. *Design and Procedure*

In a within-subject design all participants completed a single 'control' reading condition with a learning and recognition phase. The learning phase was identical to that of Experiment 1. During the recognition phase, the 80 recognition list items were presented on screen for 2000ms. The participant was then required to read the item aloud and their answers were noted by hand and entered into the computer at a later stage. The next screen required participants to estimate confidence in their decision with a number between 1 and 5, where 1 represented a complete guess and 5

represented total confidence. This screen was then replaced by a drift-correction dot in the centre of the screen before presentation of the next item. Confidence estimates were recorded on the computer after each recognition item.

3.2.1.4. *Pupil Recording*

As for Experiment 1.

3.2.2. Results

3.2.2.1. *Behavioural Data*

Participants performed at ceiling, correctly reading 100% of old and new items with maximum confidence (see Figure 3-6).

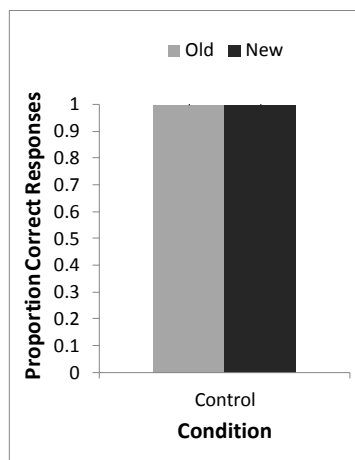


Figure 3-6: Proportion of correct responses in each condition. Error bars show standard error of mean.

3.2.2.2. *Pupil-Size Data*

Average PDR for old and new items was calculated for the control reading condition.

As PDR is a function of baseline pupil-size, baseline pupil-sizes to old and new items

were compared to ensure that any differences in PDR were not due to baseline differences. The difference was not significant ($t(24) = 1.46, p > .05, ns, r = .082$).

There was no PONE in the 'control' reading experiment, PDR for old ($M = 1.119, SD = .0535$) and new items ($M = 1.127, SD = 0.0628$) were not significantly different ($t(24) = 1.44, p > .05, ns$; see Figure 3-7). All items were included in the analysis as none were read incorrectly.

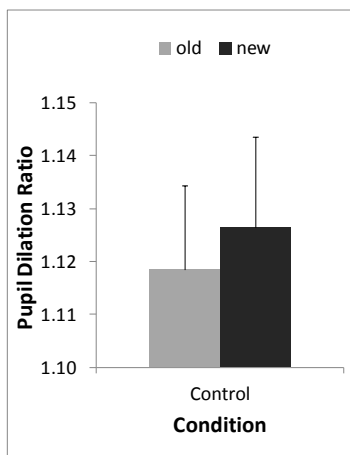


Figure 3-7: Pupil dilation ratio for old and new items in each condition. Error bars show standard error of mean.

3.2.3. Discussion

Together, Experiments 1 and 1b sought to replicate the relative increase in pupil-size that occurs when participants view previously learned items during a recognition test compared to novel items, and to determine whether it also occurs when stimuli are presented too briefly to evoke conscious recollection by participants, but may nonetheless reveal effects of prior learning.

In Experiment 1 participants' pupil-sizes increased to a greater extent when they viewed old items compared to novel items in the explicit condition, a replication of the PONE found by previous researchers. Items that had been deeply encoded

(semantic) or shallowly encoded (surface) produced a larger pupil-size than new items, in the explicit condition, and semantic items were larger than surface items (at trend level), similar to the findings of Otero et al. (2011).

If, as Võ et al. (2008) argue, pupil-size reflects cognitive effort, the PONE in the explicit condition should be smaller for deeply encoded items (less effort needed recollection). However, it was not, and the finding that the PONE is larger (at trend level) for deeply encoded items supports Otero et al.'s (2011) suggestion that pupil-size reflects memory “strength” – the PONE is larger for deeply encoded items because they evoke a stronger memory. The LOP manipulation influenced the behavioural data: in the explicit condition, more semantic items were correctly identified than surface items, whereas in the implicit condition there was no effect of LOP on performance.

Importantly, the standard PONE was not present in the implicit condition, where pupil-size was larger for new items compared to old items. This pattern of results is similar to those of Laeng et al. (2007) who looked at implicit memory in amnesic patients and found larger pupil-sizes to novel stimuli. In the absence of the PONE, a “novelty” response may have been visible instead, in the form of a larger pupil to novel stimuli than non-novel stimuli (Laeng et al., 2007; Lynn, 1966; Pavlov, 1927; Sokolov, 1963). An alternative explanation might be linked to presentation duration – the increased difficulty of the task of reading novel stimuli only presented for 33ms compared to learned stimuli that had been primed and would be easier to read even at brief duration. The implicit condition had larger PDRs compared to the explicit condition and the control reading task in Experiment 1b. This finding might be explained by an element of increased cognitive effort in that the overall difficulty of the task has been increased by the decreased presentation duration (see Chapter 1, section 1.2.2.1).

Experiment 1b was designed to explore the impact of task demands on pupillary responses, by asking participants to read the stimuli rather than make an old/new judgement. Interestingly, no PONE was found, suggesting that the PONE may occur as a result of the requirement to make a recognition decision, rather than as an automatic process resulting from the presentation of learned stimuli.

It is probably to be expected that trials that lead to a high level of confidence were the same trials that had a “strong” memory and therefore a larger PDR. However, although participants were more confident in giving old responses to old items than new responses to new items, when only considering trials with a high confidence rating (4 or 5), average PDR to correctly identified old items was still significantly larger than for correctly identified new items.

Together, the results of Experiments 1 and 1b suggest that the increase in pupil-size that occurs when participants encounter previously studied items, and recognise them as old, reflects neurocognitive processes associated with explicit, but not implicit recognition memory, and that the pupillary response is a function of task demands (recognition memory test) as opposed to being the automatic consequence of being exposed to items previously encountered during a learning phase (as in the control reading condition), or an artefact of level of confidence.

There were a number of methodological issues that limit the extent to which further inferences can be made. Whilst Experiment 1b was intended as a control task, it differed to both the explicit and implicit conditions of Experiment 1 on both task requirements and presentation duration, and therefore did not help to explain the results of the implicit condition. Although there was an LOP effect, whereby a higher number of semantically processed items were recalled than surface items in the explicit condition, the implicit condition was too easy – participants had no difficulty

reading nearly all of the stimuli during the perceptual recognition test – old or new, and as such it was not possible to tell whether more surface items than deep items were also recalled in the implicit condition, as would be predicted by perceptual fluency. Many participants also reported verbally that they had become aware that the implicit condition was a memory test for the items they had just learned, so it is unlikely that the short presentation time was a truly “implicit” test of memory. These and other methodological issues are addressed in Experiment 2.

3.3. Experiment 2 – Short vs. Long Presentation Duration

One of the key issues with the design of Experiment 1 is that it did not allow the effects of presentation duration (2000ms vs. 33ms) and task (reading vs. recognition) on pupil-size to be separated. It was not clear whether the absence of the PONE in the implicit condition arose because the manipulation had allowed old items to be read more easily due to priming (implicit test of memory), or because participants had to read stimuli rather than make an old/new judgement on them. In Experiment 2 these confounds were removed and the design strengthened by adopting a within-subject approach and replacing the concepts of “explicit” and “implicit” tests of memory with 2000ms exposure (long duration) and visual perceptual threshold exposure (short duration) during the recognition phase, and for each exposure duration asking participants to either read or identify the word as old or new.

Another methodological issue that arose in Experiment 1 was that in the “implicit” condition most participants perceived all stimuli very clearly, whereas some couldn’t read any of the stimuli at all. As a result, in Experiment 2, a thresholding program was used to calculate individual presentation durations for each participant, by increasing or decreasing presentation duration until approximately 60% of short duration items could be correctly identified. Within the implicit literature, the

Cheesman and Merikle (1984) distinction differentiates an objective threshold for presentation duration at which participants perform at chance because they are genuinely guessing, and a slightly higher subjective threshold when presentation duration produces the feeling of guessing at point of recognition, however participants perform at levels above chance, but without consciously recollecting stimuli. This type of marginally perceptible or “subliminal” (Cheesman, & Merikle, 1984) presentation allows memory to be tested “implicitly” (e.g., Chan, 1992; Cheesman, & Merikle, 1984; 1986; Dienes, Altmann, Kwan, & Goode, 1995; Dienes, & Berry, 1997; Merikle, 1992).

A further methodological improvement was the introduction of an isoluminant visual mask. Unless followed by a mask, briefly presented stimuli can leave an “afterimage” created by temporary pigment changes in the photoreceptors of the retina, which result in negative images of the stimuli persevering beyond the brief presentation. Experiment 2 included a mask of 6 ampersands (“&&&&&”) both before and after stimuli in the same size and font. The mask also minimised any change in screen luminance from a blank screen to one showing a stimulus (see Chapter 2, section 2.1.2.2). In order to reduce overall accuracy levels, items appeared either above or below a fixation cross at random. This served to make it more difficult to read the stimulus because it was not fixated and its position could not be predicted; it was necessary because pilot testing of the thresholding program revealed that even at the minimum presentation duration of a single monitor refresh (10ms at 100Hz), centrally positioned stimuli could be read by most participants. Finally, in order to simplify the design, the LOP manipulation and confidence measure were removed, allowing the effects of task and duration manipulations to be seen more clearly.

Following on from the results of Experiments 1 and 1b, the PONE was predicted to be present in the long-duration recognition memory condition (as per the explicit condition), absent for long duration reading conditions (no difference in maximum pupil-size for correctly identified new and old items as per the control condition), and reversed for the short-duration condition (like the implicit condition). No prediction was made for the short-duration recognition condition as it was not known whether it was duration or task requirements impacting on pupil-size in Experiments 1 and 1b), however if participants are able to make old/new judgements accurately at short-duration presentations, then the PONE may be observed. Consistent with Experiment 1, due to perceptual fluency, more correct old responses than new responses were expected in the short duration reading condition.

3.3.1. Method

3.3.1.1. Participants

Twenty-eight participants (2 male; age range: 18.3-49.4, $M = 23.1$, $SD = 7.19$), with normal or corrected-to-normal vision were recruited from the psychology course-credit and subject pools at the University of Sussex, and through personal contact.

Participants were briefed with a detailed consent form and verbal description of the experiment, and invited to ask questions. Written consent was obtained prior to testing and participants were fully debriefed at the end. The experiment was approved by the relevant ethics committee.

3.3.1.2. Materials/Apparatus

Using nouns selected from the MRC Psycholinguistic Database nine word lists were created. Four learning lists each comprised 40 items, and the four recognition lists contained the 40 nouns from the corresponding learning list plus 40 new items. An

extra 40 item list was created to determine individual thresholds for the short duration conditions. All words were 6 letters long, and lists were matched for concreteness (learning items range = 259-646, $M = 486$; new items range = 254-652, $M = 489$; calibration items range = 267-643, $M = 487$), familiarity (learning items range = 277-628, $M = 509$; new items range = 256-632, $M = 512$; calibration items range = 269-634, $M = 511$), imageability (learning items range = 280-643, $M = 506$; new items range = 289-643, $M = 511$; calibration items range = 301-646, $M = 512$) and frequency (learning items range = 21-478, $M = 54$; new items range = 18-472, $M = 56$; calibration items range = 20-483, $M = 54$), according to the K-F norms.

The learning and recognition lists were presented in black 20pt Arial font in the centre of a light grey background under fixed illumination. Items were presented using Experiment Builder software (SR-Research, Ontario). All items are presented in Appendix B. During the recognition test, pupil-size was recorded using an EyeLink 1000 tower-mounted eye-tracker with a temporal resolution of 2ms and a spatial resolution of around 0.15 degrees. The stimuli were displayed on a 21 inch CRT monitor with a screen resolution of 1,280 x 1,024 pixels and a refresh rate of 100Hz. Actual screen dimensions were 40cm horizontal and 30cm vertical. Participants were seated with their chin on a rest 70cm from the screen in an adjustable chair that had been modified to prevent any rotational movement.

3.3.1.3. *Design and Procedure*

The experiment comprised four conditions, two recognition memory conditions, one short and one long-duration, and two reading conditions, one short and one long-duration. In a within-subject design all participants completed all four conditions. Before the main experiment started, participants completed a thresholding task using the extra 40 noun list. List items replaced a mask (“&&&&&”) either above or below

the fixation cross. This procedure was used because Experiment 1 showed that without this unpredictability in location most people could read the items, even at the lowest duration when looking directly at a single mask. Starting with a presentation time of 100ms, the program shortened presentation duration by 10ms when participants read an item correctly and lengthened it by 10ms when they failed to identify the stimulus. After 40 items had been presented the proportion of correct responses at the various presentation durations was displayed; the experimenter was then able to select a presentation duration that was neither so short that participants were unable to read the majority of items, nor so long that they were performing at ceiling. The duration chosen for most participants was the shortest duration that resulted in an equivalent number of correct and incorrect responses. Where there was no duration at which numbers were equal, the duration at which correct responses were greater than incorrect responses was chosen.

All conditions contained a learning and recognition phase. The learning phase was identical across the conditions. The 40 learning list items were presented in the centre of the screen for 2000ms and participants were asked to try to remember them. For each recognition phase, list items replaced a 500ms mask either above or below a central fixation cross. The order in which the four conditions were performed was rotated across participants. For the long duration reading condition the 80 recognition list items were presented and left on screen for 2000ms. The participant was asked to read them out loud as they were presented. This screen was then replaced by a drift-correction dot in the centre of the screen before presentation of the next item. The procedure was identical for the long duration recognition condition but participants were instead asked to state whether the item was old (previously encountered in the learning phase) or new (not previously encountered). For the short duration reading condition the 80 recognition list items replaced the mask for the

brief length of time determined in the thresholding task for that participant. Items were then remasked for 2000ms. Participants were asked to read the words out loud (or give their best guess) as they were presented. This screen was then replaced by a drift-correction dot in the centre of the screen before presentation of the next item. For the short duration recognition condition the procedure was identical except participants were asked to state whether the item was old (previously encountered in the learning phase) or new (not previously encountered). In order that participants were looking at the item during its brief presentation, in the short duration conditions participants were advised to blink only whilst the drift-correct dot was on screen. In all conditions, old/new judgements and correct/incorrect reading responses were recorded on the computer after each item.

3.3.1.4. Pupil Recording

Maximum pupil-size was recorded from the left eye during each recognition period – the time for which the item was on screen for the long duration conditions, and the time for which the item and the mask were on screen for the short duration conditions. A Pupil Dilation Ratio (PDR; see Chapter 2, section 2.1.2.1) was calculated expressing the maximum pupil-size for each 2000ms recognition trial as a proportion of the maximum pupil-size during that trial's 250ms baseline.

3.3.2. Results

3.3.2.1. Behavioural Data

The proportion of correct responses to old and new items was calculated for each condition. In all four conditions participants performed significantly above chance (50%) at correctly identifying old and new items (all $t_s > 2$, $p_s \leq .05$), demonstrating that the threshold measuring task worked – participants were not performing at ceiling

or floor levels (with the exception of the long reading condition where participants achieved 100% correct responses). A 2 x 2 x 2 ANOVA on proportion of correct responses, with within-subject factors of task (reading vs. recognition), presentation duration (long vs. short), and item-type (old vs. new), showed a significant main effect of presentation duration ($F(1,27) = 64.3$, $MSE = 0.035$, $p < .001$, $\eta_p^2 = .700$) – participants made more correct responses at long durations compared to short durations. The main effect of task was also significant ($F(1,27) = 91.1$, $MSE = 0.032$, $p < .001$, $\eta_p^2 = .770$) – participants made more correct responses in the reading conditions compared to the recognition conditions. The main effect of item-type was not significant ($F(1,27) = 1.98$, $p > .05$, ns; see Figure 3-8).

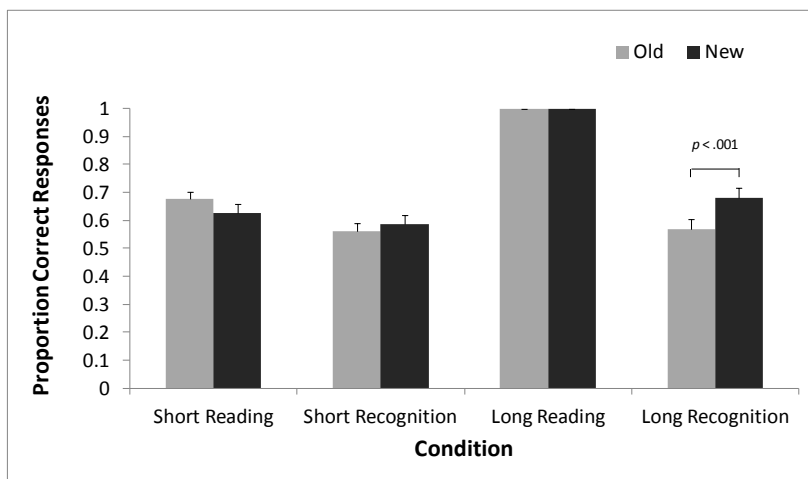


Figure 3-8: Proportion of correct responses for old and new items in each condition. Error bars show standard error of mean.

These main effects were qualified by a number of interactions, including a significant task by presentation duration interaction ($F(1,27) = 61.9$, $MSE = 0.020$, $p < .001$, $\eta_p^2 = .700$) – increased presentation duration produced a larger improvement in participants' correct responses in the reading task (34.8%, $SD = 13.9\%$) than in the recognition task (5.04%, $SD = 18.8\%$; $t(27) = 7.86$, $p < .001$, $r = .696$). The interaction between task and item-type was also significant ($F(1,27) = 11.0$, $MSE = 0.011$, p

$<.01$, $\eta_p^2 = .290$) – participants responded correctly to more old items (83.8%, $SD = 7.36\%$) than new items (81.4%, $SD = 8.05\%$) in the reading tasks ($t(27) = 1.90$, $p = .068$, $r = .118$), but to more new items (63.3%, $SD = 14.7\%$) than old items (56.6%, $SD = 12.7\%$) in the recognition tasks ($t(27) = 2.56$, $p < .01$, $r = .195$). Finally, the presentation duration by item-type interaction also reached significance ($F(1,27) = 10.9$, $MSE = 0.006$, $p < .01$, $\eta_p^2 = .291$) – participants responded correctly to old items more often than new items at short duration ($M = 1.25\%$, $SD = 0.13\%$), and to new items more often than old items at long duration (5.58%, $SD = 5.0\%$; $t(27) = 3.30$, $p < .01$, $r = .287$). The three-way interaction was not significant ($F(1,27) = 1.64$, $p > .05$ ns), however *a priori* t-tests showed that for the short reading condition, participants were able to correctly read more old items (67.6%, $SD = 14.7\%$) than new items at trend level (62.8%, $SD = 16.1\%$; $t(27) = 1.90$, $p = .07$, $r = .118$). Unlike the long duration recognition condition (and the ‘explicit’ condition of Experiment 1), where more new items were correctly identified than old items ($t(27) = 4.34$, $p < .001$, $r = .411$), in the short duration recognition condition correct identification of old and new items was very similar ($t < 1$, ns).

3.3.2.2. Pupil-Size Data

Average PDRs for old and new items were calculated for all conditions. As PDR is a function of baseline pupil-size, baseline pupil-sizes for old and new items in each condition were compared to ensure that differences in PDR were not due to baseline differences. The difference was not significant ($F(1,27) = 1.21$, $p > .05$, $\eta_p^2 = .043$).

A 2 x 2 x 2 ANOVA on PDR, with within-subject factors of task (reading vs. recognition), presentation duration (long vs. short), and item-type (old vs. new) revealed a significant main effect of presentation duration ($F(1,27) = 19.0$, $MSE = 0.002$, $p < .001$, $\eta_p^2 = .413$) – PDR values were greater when items were presented for

short durations. There was no main effect of item-type ($F(1,27) = 0.224, p > .05, ns$), and no significant main effect of task ($F(1,27) = 0.152, p > .05, ns$). The main effect of presentation duration was qualified by a significant interaction with task ($F(1,27) = 6.72, MSE = 0.003, p < .01, \eta_p^2 = .199$) – there was a trend for stimuli displayed for a short duration to produce a larger PDR in the reading task ($M = 1.148, SD = 0.059$) than the recognition task ($M = 1.134, SD = 0.059, t(27) = 1.93, p = .06, r = .122$), whereas stimuli displayed for a long duration produced a larger PDR in the recognition task ($M = 1.123, SD = 0.053$) than in the reading task ($M = 1.103, SD = 0.045, t(27) = -2.14, p < .05, r = .145$). There was no significant interaction between duration and item-type ($F(1,27) = 0.863, p > .05, ns$) or between task and item-type ($F(1,27) = 2.93, p = .098, ns$), and the three-way interaction was not significant ($F(1,27) = 0.738, p > .05, ns$; see Figure 3-9).

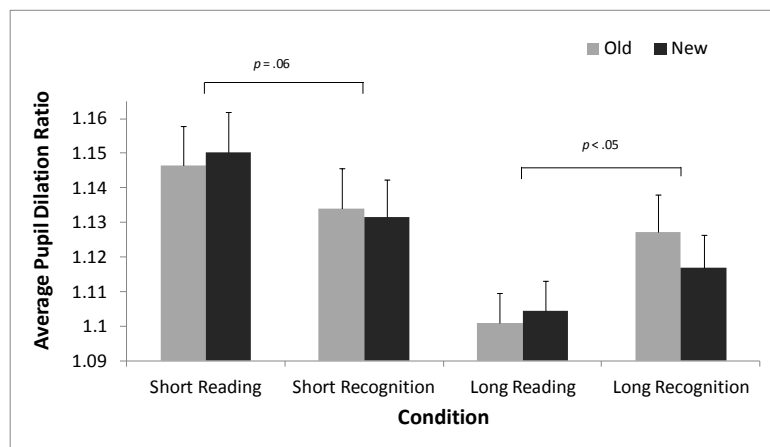


Figure 3-9: Pupil dilation ratio for old and new items in all four conditions. Error bars show standard error of mean.

The analysis was repeated with PDRs calculated across only those trials in which old and new items were correctly identified. A 2 x 2 x 2 ANOVA on PDR, with within-subject factors of task (reading vs. recognition), presentation duration (long vs. short), and item-type (old vs. new) now revealed a significant main effect of item-type ($F(1,27) = 14.3, MSE = 0.001, p < .001, \eta_p^2 = .352$) – average PDR was greater for old

items compared to new items. The main effect of presentation duration was also significant ($F(1,27) = 18.8$, $MSE = 0.003$, $p < .001$, $\eta_p^2 = .410$) – PDR values were greater when items were presented for short durations than long durations. There was no significant main effect of task ($F(1,27) = 0.20$, $p > .05$, ns). These main effects were qualified by a significant task by presentation duration interaction ($F(1,27) = 6.71$, $MSE = 0.002$, $p < .01$, $\eta_p^2 = .203$) – like before, stimuli displayed for a short duration produced a larger PDR in the reading task ($M = 1.150$, $SD = 0.058$) than the recognition task at trend level ($M = 1.135$, $SD = 0.061$, $t(27) = 1.78$, $p = .086$, $r = .105$), whereas stimuli displayed for a long duration produced a larger PDR in the recognition task ($M = 1.123$, $SD = 0.052$) than in the reading task ($M = 1.103$, $SD = 0.045$, $t(27) = -2.06$, $p < .05$, $r = .136$). There was also a significant interaction between task and item-type ($F(1,27) = 21.11$, $MSE < 0.001$, $p < .001$, $\eta_p^2 = .44$) – PDR for correctly identified old items ($M = 1.141$, $SD = 0.056$) was significantly larger than for correctly identified new items ($M = 1.116$, $SD = 0.046$) in the recognition tasks ($t(27) = 5.31$, $p < .001$, $r = .510$); however there was little difference in PDR for correctly identified old ($M = 1.120$, $SD = 0.047$) and new items ($M = 1.122$, $SD = 0.047$) in the reading tasks ($t(27) = -0.361$, $p > .05$, ns) – in other words there was a pupil old/new effect in the recognition conditions but not in the reading conditions. There was no significant interaction between duration and item-type ($F(1,27) = 0.31$, $p > .05$, ns) and the three-way interaction was not significant ($F(1,27) = 1.17$, $p > .05$, ns; see Figure 3-10).

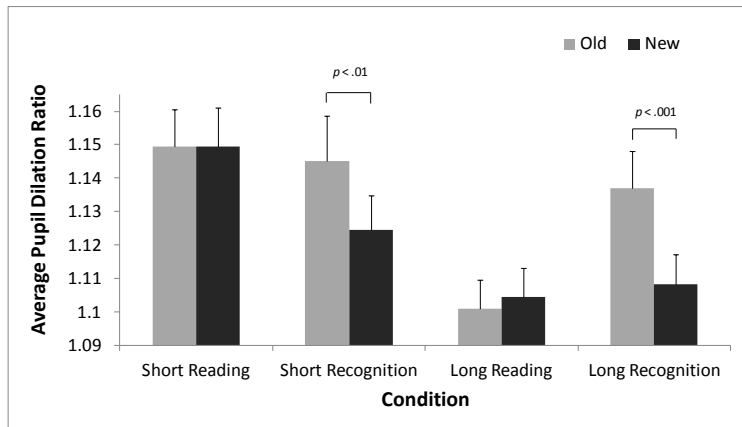


Figure 3-10: Pupil dilation ratio for correctly identified old and new items in all four conditions. Error bars show standard error of mean.

3.3.3. Discussion

Experiment 2 systematically contrasted the effects of presentation duration (short vs. long) and task (reading vs. recognition) on pupil-size when participants were presented with old and new items. As predicted, in the long duration recognition condition participants' maximum pupil-size was significantly larger for correctly identified old items compared to correctly identified new items. Interestingly, the PONE was also present at short duration presentations, even though recognition performance was significantly worse. This old/new effect was not observed for the reading conditions, either at short or long duration. Taken together these findings suggest that the reversed effect in the implicit condition of Experiment 1 was most likely due to differences in task demands (reading vs. recognition), not presentation duration – the PONE occurs whenever participants are asked to make a recognition judgement on a word, even when presented very briefly, but is not present when participants are asked to read a word out loud without making a recognition judgment, even when that word is present for a long duration.

Although recognition performance was worse than performance on the reading task, recognition rates for old items were comparable across short and long durations, but

recognition rates for new items were impaired in the short duration compared to the long duration recognition condition. However, overall rates of correctly recognising old items as old were poor with 6 participants performing below chance in the short recognition condition and, perhaps more surprisingly, 14 participants in the long recognition condition, 3 of whom performed poorly in both recognition conditions. There are many factors affecting recognition memory, including age, attention and context. Although condition order was rotated across participants, one possible explanation of the poor recognition performance in the present experiment could be a build up of retroactive interference across word lists, as by the fourth condition participants had been asked to remember a large number of items. This is a disadvantage of the within-participants design employed.

Another interesting finding to emerge from Experiment 2 was that overall PDR was greater when stimuli were presented for a short duration. A possible explanation for this finding might be that the short durations increased the cognitive effort required to perform the recognition/reading tasks. Increased “cognitive load” could also account for the larger pupil-sizes seen in the implicit condition of Experiment 1, when stimuli were presented for a short time. However it is important to note that the PONE is not overwhelmed by effort-related increases in pupil-size due to increased task difficulty because it is still present in the short duration recognition task.

3.4. General Discussion

The Experiments in Chapter 3 sought to replicate the PONE – the relative increase in pupil-size that occurs when participants view previously learned items during a recognition test compared to novel items, and to build on Otero et al.’s (2011) finding of a pupil-size difference between recollection and familiarity ratings, by asking

whether the PONE still occurs when stimuli are presented too briefly to evoke conscious recollection, but may nonetheless reveal effects of prior learning.

The evidence from Experiments 1 and 2 replicates the PONE demonstrated by previous research (see Chapter 1, section 1.3.2.3), and shows that maximum pupil-size is larger when participants encounter previously studied items, compared to new items, when carrying out an explicit recognition memory task. Importantly Experiment 2 extends this finding to show that this pupillary old/new effect is a function of task demands (recognition memory), as opposed to being an automatic consequence of being exposed to items previously encountered during a learning phase (as in the reading conditions), and also occurs for stimuli presented for brief durations. The problems identified in Experiment 1, such as participants performing at ceiling in the implicit condition, were addressed in Experiment 2. A better comparison of task demands at short and long durations was allowed by the introduction of a short duration reading condition.

The LOP manipulation in Experiment 1 had the expected effect on performance, increasing successful recognition for deeply encoded (semantic) items compared to shallowly encoded (surface) items in the explicit condition, but having no effect on performance in the implicit condition. When old responses were collapsed over surface and semantic items, more new than old items were correctly identified in the explicit condition, and more old than new items were correctly identified in the implicit condition, suggesting that perceptual fluency may enhance recognition performance (e.g., Jacoby, & Dallas, 1981). This was also the case in Experiment 2.

The results of the LOP manipulation suggest that, contrary to Võ et al.'s (2008) cognitive effort explanation of the PONE, items which had been deeply encoded produced a larger pupil-size than shallowly encoded items (albeit at trend level),

which in turn produced a larger pupil-size than new items in the explicit condition.

This pattern of results is similar to Otero et al.'s (2011) findings of a graded pupil-size for LOP and supports their suggestion that instead of a difference in cognitive effort between old and new items, pupil-size may reflect memory "strength" – the PONE is larger for deeply encoded items because they evoke a stronger memory.

Interestingly, rather than the standard PONE, the implicit condition revealed the reverse – larger pupil-size for new items compared to old items, a finding that is similar to those of Laeng et al. (2007) who looked at implicit memory in amnesic patients and found larger pupil-sizes to novel stimuli. The difference was driven by surface items, which had the smallest pupil-size, with semantic items intermediate but not significantly different to either new or surface items. A possible explanation of this pattern of results is that a pupil orienting response to novel items (e.g., Lynn, 1966; Nieuwenhuis et al., 2011a; Pavlov, 1927; Sokolov, 1963) was no longer masked by the PONE when presentation was too brief to elicit conscious recollection of stimuli. An alternative explanation is that the increased pupil-size to new items reflects the increased difficulty of the short duration, relative to old items which were made easier to read through perceptual fluency. As it was difficult to tell whether task demands or presentation duration lead to the results in the implicit condition, a control reading task was carried out with the duration of the explicit condition, but where participants were simply required to read the items out loud as in the implicit condition. In this task, performance and pupil-size was equal for old and new items, and it was not possible to draw any conclusions. Experiment 2 was designed as a more complete orthogonal comparison which allowed the effects of presentation duration (2000ms vs. 33ms) and task (reading vs. recognition) to be separated.

The key finding from Experiment 2 was that the PONE was present in both duration recognition conditions, but was absent from both duration reading conditions.

Therefore the PONE occurs whenever participants are asked to make a conscious recognition judgement on a word, even when presented very briefly, but is not present when participants are asked to read a word without making a recognition judgment, even when that word is present for a long duration. These findings may provide more evidence in support of Otero et al.'s (2011) memory strength explanation of the PONE, rather than Võ et al.'s (2008) cognitive effort explanation, because the PONE is not overwhelmed by effort-related increases in pupil-size due to increased task difficulty in the short conditions – there were separate main effects of presentation duration (effort) and item-type (memory strength) on PDR, with no duration by item-type interaction. The decrease in presentation duration also had more of an effect on pupil-size in the reading conditions than in the recognition conditions.

The results of Experiment 2 suggest that the reversed pupil effect in the implicit condition of Experiment 1 was most likely due to task demands (reading), rather than presentation duration, but it is unclear why this was not replicated in the short reading condition of Experiment 2. An explanation may lie in part in the fact that the presentation duration was customised for participants in Experiment 2 such that they achieved ~60% correct identification of items, whereas in Experiment 1 participants were able to correctly identify 89% of items. Taken together, Experiments 1 and 2 suggest that the changes in pupil-size that occur, when participants encounter previously studied items, reflect neurocognitive processes associated with making a recognition judgement that do not occur when simply reading, even within the context of a 'memory experiment'. Previous research has shown that pupil-size is affected by explicitly making a decision in other types of study (see Chapter 1, section 1.2.3.6).

However at short duration the PONE was only apparent when analysis is limited to correctly identified items. In this condition, it was not possible to determine which items were presented at the subjective threshold of conscious awareness (Cheesman, & Merikle, 1984) and their relationship with pupil-size. It is also not possible to exclude the possibility that the perceptual recognition task was contaminated with recollective processes in the short reading condition, with some participants anecdotally reporting conscious recognition of stimuli.

To address these issues, Chapter 4 adopts a widely used implicit learning procedure in order to establish whether the PONE occurs when participants encounter letter strings that either do or do not conform to a learned artificial grammar. The advantage of this approach is that whilst for half of the strings, the letter order follows the same artificial grammar rules as the strings encountered in the learning phase, at recognition all letter strings are novel and have not been previously encountered.

4. The Role of Conscious Awareness in the PONE

Implicit Grammar

As argued in the discussion of Experiments 1 and 2, limiting presentation to a brief duration during recognition was not particularly successful for determining whether the PONE is also present when memory is tested implicitly. Stimuli in Experiment 2 were presented for a duration derived for each participant individually as the number of monitor refreshes at which ~60% of practice items could be read. It was hoped that this duration would produce a significant proportion of words at the subjective threshold (Cheesman, & Merikle, 1984), eliciting the feeling of guessing at the actual point of recognition, but a performance significantly above chance. However, as some participants reported explicit awareness of some of the recognition items, it was not possible to draw conclusions about implicit recognition.

There is debate in the literature regarding the criteria for “implicitness”. Shanks and St John (1994) argue that knowledge elicited by cued recall and forced choice tests is not implicit, given that it is far too difficult to be sure that you have excluded explicit influences on a task that implicitly tests explicitly learned knowledge. A more “process pure” method than implicit tests of memory, which researchers have employed, is to look at implicit learning using paradigms such as Artificial Grammar Learning (AGL). The literature demonstrates that implicit learning is also very complex and difficult to define simply as, for example, learning without conscious awareness, because consciousness and awareness are also very complex and subjective concepts that are problematic to define or measure (Cleeremans, & Dienes, 2008), for example, Frensch (1998) describes eleven different definitions of implicit learning.

Berry and Dienes (1993) argue that implicit learning is unintentional, happening without conscious awareness, and that the resulting knowledge is relatively inaccessible with free report. Explicit learning, in contrast, is usually hypothesis-driven (an attempt to define, test and refine rules or concepts whilst learning) and happens with conscious awareness of both the learning experience and the knowledge acquired (see also Cleeremans, & Dienes, 2008; Dienes, & Berry, 1997).

Research has shown that implicit learning is associated with attention to stimuli rather than on underlying rules, and is more robust to neuropsychological impairment (e.g., Knopman, & Nissen, 1987; Nissen, & Bullemer, 1987; Nissen, Willingham, & Hartman, 1989; Schacter et al., 1988). Knowlton, Ramus and Squire (1992) found that amnesic participants were able to classify new items as well as control participants (63% and 67%, respectively), but were less able than controls to correctly recognise old items (62% and 72%, respectively). Individual differences, such as age and IQ, have less of an effect on implicit learning than explicit learning (e.g., Cherry, & Stadler, 1995; Frensch, & Miner, 1994; Howard, & Howard, 1989; 1992; Myers, & Connor, 1992). Reber, Walkenfeld and Hernstadt (1991) showed that AGL performance was less correlated with IQ than performance on a problem-solving task.

The first AGL experiments used to investigate implicit learning were carried out in the 1960s by Reber (1967; 1969; 1989). Participants were shown a series of letter strings that obeyed a 5-letter finite-state artificial grammar (a network with a finite number of rules/paths that can be followed from entry to exit, see Figure 4-1). They were then asked to identify 'grammatical' from 'non-grammatical' strings in a 'recognition' phase. Importantly, the grammatical strings in the recognition phase were not the same strings that participants were exposed to in the learning phase – they simply followed the same artificial grammar rules. Participants correctly

identified 69% of the previously unseen grammatical strings, which Reber (1967) proposed as evidence of implicit learning of the underlying grammatical rules. Reber reported that participants emerged with some knowledge but were not able to explain it fully. This was more formally tested by Dienes, Broadbent and Berry (1991) who asked participants to describe rules or strategies that could be used by someone who had not seen the stimuli. Using those rules three independent judges rated each string as grammatical or ungrammatical to simulate a classification performance of 54%. This result was significantly less than actual classification performance (65%) indicating that participants' free-report of learned knowledge was impoverished compared to their ability to apply it.

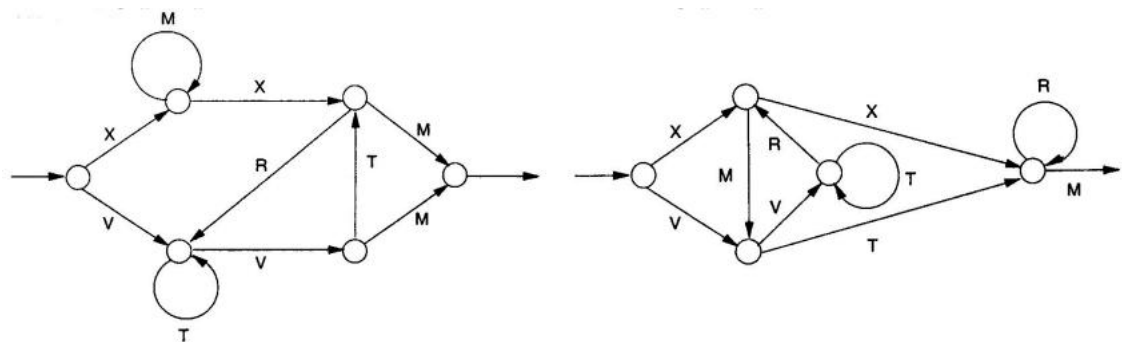


Figure 4-1: The two finite state grammars used to generate the strings (from Reber, 1969).

Since Reber's original experiments, AGL has become a widely used technique for investigating implicit learning, and his findings have been widely replicated. Several different accounts have been proposed for the basis of above chance classification judgements including: abstracting rules about the underlying grammar; memorising whole strings or fragments of the training stimuli (chunking; Wickelgren, 1979); and learning the statistical relationships between fragments of the strings such as in connectionist models (see Pothos, 2007, for a comprehensive review).

Dienes and Berry (1997) argue that the learning of underlying correct classification of grammatical strings in AGL experiments is implicit because, unless participants are

instructed to look for rules, learning occurs without intent or conscious awareness, and participants perform at above chance whilst not being able to verbalise how they are making their decisions. Confidence in grammatical judgements is also often unrelated to performance, indicating an absence of metaknowledge (Chan, 1992; Dienes et al., 1995; Dienes, & Altman, 1997).

Previous research and connectionist modelling have indirectly suggested that subjective feelings of familiarity may be central to implicit learning (e.g. Norman, Price, Duff, & Mentzoni, 2007; for a review see Cleeremans, & Dienes, 2008). However, there had been no direct tests of this hypothesis; Scott and Dienes (2008) provided the first direct evidence that structural similarities in grammatical as opposed to non-grammatical letter strings (such as fragment frequency and repetition structure) are experienced subjectively as familiarity, by examining how grammatical judgements and confidence related to participants' reports of familiarity. In a series of AGL experiments they asked participants to state whether each test string was grammatical, how certain they were in their decision, how they made their decision, and how familiar the string felt to them. Participants were able to correctly identify around 60% of the grammatical strings – significantly above chance. Importantly, participants' subjective familiarity ratings for strings predicted their grammaticality judgements and the degree to which they were correct in their response – the more familiar a string felt, the more likely participants were to rate it as grammatical, and the more likely they were to be correct in their judgement, even in the absence of confidence in their response (Scott, & Dienes, 2008).

As discussed in section 1.3.1.2 of the introductory chapter, some current models of recognition memory argue that studied items can be recognized as old on the basis of two separate processes – a conscious recollective process that involves retrieval of

specific contextual information concerning the original learning experience (sometimes represented by a “remember” or “R” response) and a familiarity based process that provides a context-free sense that an item has been previously encountered (a “know” or “K” response; e.g. Yonelinas, & Jacoby, 1996). Many studies show that the two processes can be differentiated behaviourally and these findings are used to support the argument that they have different underlying neural mechanisms (see Yonelinas, 2002, for a review). A further line of evidence supporting dual-process models is the ERP old/new effect in which specific neuroelectric signals appear to be linked to R/K responses (e.g., Curran, 2004; Duarte et al., 2004; Düzel et al., 1997; Rugg, Schloerscheidt, & Mark, 1998; Trott, Friedman, Ritter, Fabiani, & Snodgrass, 1999; Wolk et al., 2006; see Chapter 6, section 6.1.1).

An obvious question is therefore whether similar differentiation occurs for the PONE – in other words does it differ in magnitude between items that are consciously recollected compared to those that are recognised on the basis of familiarity alone. If it does, then any explanation of the PONE in recognition memory must therefore take account of both recollection and familiarity processes and the ways in which they dissociate experimentally. Previous research into the PONE (see Chapter 1 section 1.3.2.3 and Chapter 3 section 3.4) has not yet answered this question. Using a remember-know paradigm Otero et al. (2011) found that the PONE was larger when participants responded “remember” (recollection), compared to when they responded “know” (familiarity). Critically, the PDR for “know” items was greater than the PDR when participants viewed novel items (although this difference was only significant at a trend level). Otero et al.’s (2011) suggestion that the PONE reflects an aggregate strength of memory signal, based on both familiarity and recollective processes, was prompted by recent accounts of recognition memory that reject the idea that

recollection, when it occurs, is 'all or nothing', and that familiarity processes are only important in the absence of recollection (Wixted, 2007a; Wixted, & Stretch, 2004). Instead these accounts suggest that both recollective and familiarity signals vary on a continuum, and items are judged to be old when an aggregate signal exceeds a certain threshold. Items that are recognized on the basis of familiarity, in the absence of conscious recollection, would still have a greater overall strength of memory signal than new items, but the strength of memory signal would be considerably weaker than the signal associated with items that are recollected.

Experiment 3 aims to replicate the previous findings of increased PDR for old compared to new words in an explicit recognition condition. Additionally, Experiment 3 aims to further investigate the mnemonic processes associated with pupil dilation by measuring pupil-size in an implicit recognition condition using AGL, proposing that for the implicit condition, conscious recollection will not be available as both the "grammatical" and "nongrammatical" strings presented in the recognition phase will be different to those presented in the learning phase. Previous research (Reber, 1967; 1969; Scott, & Dienes, 2008; 2009) has indicated that implicit learning of grammatical letter strings evokes a greater sense of familiarity than non-grammatical. This would suggest that familiarity alone is enough to elicit a memory signal exceeding recognition threshold. It was predicted that a PONE would occur in the implicit condition, reflecting familiarity signal strength, but that this effect would be smaller than for the explicit test of memory.

4.1. Experiment 3 – Implicit Grammar

4.1.1. Method

4.1.1.1. *Participants*

Twenty-four participants (8 male; age range: 19.0-36.2, $M = 24.6$, $SD = 5.29$) with normal or corrected-to-normal vision were recruited from the psychology course-credit and subject pools at the University of Sussex, and through personal contact.

Participants were briefed with a detailed consent form (specific to the condition to which they were allocated) and verbal description, and invited to ask questions. To gain informed consent to participate, without deception or revealing the artificial grammar, participants were informed that they would be presented with letter strings in the learning phase of that test, and would have to make 'yes/no' decisions on 60 letter strings based on the strings presented in the learning phase. Participation was on a voluntary basis, and participants were thanked and debriefed with a verbal description of the aims of the study and the opportunity to ask any further questions after completing the study. The experiment was approved by the relevant ethics committee.

4.1.1.2. *Materials/Apparatus*

For the explicit condition two 30 item word lists (A and B; see appendix C) were created using nouns selected from the MRC Psycholinguistic Database. All items were 7 letters long, and lists were matched for familiarity (list A range = 420-613, $M = 521$; list B range = 442-598, $M = 524$), imageability (list A range = 274-593, $M = 425$; list B range = 258-591, $M = 410$) and frequency (list A range = 21-96, $M = 53$; list B range = 22-98, $M = 52$), according to the K-F norms. Words likely to elicit a strong emotional response were removed. Items were presented using the Experiment

Builder software associated with the EyeLink II eye-tracker (SR-Research, Ontario) in black 20pt Arial font in the centre of a light grey background under fixed illumination. During the learning phase participants saw either list A or B, and during the recognition phase participants saw all items from both lists in a randomised order.

Stimuli for the implicit condition were taken from Scott and Dienes (2008) who used two finite-state grammars from Reber (1969) to generate 45 grammatical strings, 5-9 letters long, from each grammar using the letters M, R, T, V, and X and the same set of valid starting bigrams and final letters. 15 strings from each grammar were repeated three times in a random order to create two 45 item learning lists (A and B; see appendix C). The remaining 30 items from each grammar were combined in a random order to create the recognition list. Lists were matched for string length and Scott and Dienes (2008) carried out statistics to ensure the structural similarity of learning strings to recognition strings. Items were presented using Experiment Builder (SR-Research, Ontario) in black 20pt Arial font in the centre of a light grey background under fixed illumination. During the learning phase participants saw either list A or B, and during the recognition phase participants saw new strings, 30 items generated from each grammar, in a randomised order, a total of 60 items.

Throughout the recognition phase of both conditions, pupil-size was recorded using an EyeLink II head-mounted eye-tracker with a temporal resolution of 2ms and a spatial resolution of around 0.25 degrees. The stimuli were displayed on a 21 inch CRT monitor with a screen resolution of 1,280 x 1,024 pixels and a refresh rate of 60Hz. Actual screen dimensions were 40cm horizontal and 30cm vertical. Participants were seated approximately 70cm from the screen in an adjustable chair that had been modified to prevent any rotational movement. Responses during the

recognition phase were made using the left and right triggers of a computer gamepad and recorded by the host PC.

4.1.1.3. Design and Procedure

The study employed a 2 (recognition test: explicit vs. implicit) x 2 (item-type: old vs. new) within-subjects design. All participants completed both the implicit and explicit conditions, which each contained a learning and recognition phase. Participants completed either the implicit or explicit condition first, and learned either list A or B (independently) in each condition. This gave a total of 8 different combinations (AA, AB, BB, AA x explicit first, implicit first) which were rotated across participants to avoid practice and list effects. Participants were briefed on arrival with general instructions and asked to complete a consent form. Further instructions specific to condition were displayed on a computer screen as they completed the tasks.

Explicit Condition

The experimenter selected which list (A or B) should be used for the learning phase. On-screen instructions informed participants that they would see a set of 30 words and asked them to try to remember them. The 30 learning items were presented one at a time in the centre of the screen for 2000ms. Participants were then fitted with the eye-tracker and performed a short 3 point calibration task. During the recognition phase, participants were presented with 60 items (lists A and B in a randomised order), 30 of which had been presented in the learning phase (old items), the other 30 were from the list that had not been learned (new words). At the start of each trial, a drift-correction dot was presented in the centre of the screen, followed by a mask “&&&&&&&” of 7 ampersands for 500ms, and then a recognition item for 2000ms. Masks and recognition items were presented in the same size and font as items in the

learning phase. Participants were asked to state whether the item was old (previously encountered in the learning phase) or new (not previously encountered), prompted by a screen with the word “old” on the left and “new” on the right, corresponding to the buttons on the gamepad. This screen remained visible until the participant made a decision, at which point the screen was replaced by a drift-correction dot before presentation of the next item. Old/new judgements were recorded on the computer after each recognition item.

Implicit Condition

The experimenter selected which list (A or B) should be used for the learning phase. Participants were not informed that there were rules to the letter strings during the learning phase. Onscreen instructions informed participants that they would see 45 letter strings one at a time, each on screen for 5000ms with a 5000ms gap in between to encourage attention to the unfamiliar stimuli, during which they should write down as much of the string as they could remember. No indication was given that they would be learning an artificial grammar. The eye-tracker was then fitted followed by a 3 point calibration task. During the recognition phase, participants were presented with 60 new letter strings, 30 from grammar A and 30 from grammar B. No strings in the recognition phase had been presented in the learning phase so could not be explicitly recognised. At the start of each trial, a drift-correction dot was presented in the centre of the screen, followed by a mask “&&&&&&&” of 7 ampersands for 500ms, and then a recognition item for 2000ms. Masks and recognition items were presented in the same size and font as items in the learning phase. Participants were asked to state whether the item was grammatical, prompted by a screen with the word “yes” on the left and “no” on the right, corresponding to the buttons on the gamepad. This screen remained visible until the participant made a decision, at

which point the screen was replaced by a drift-correction dot before presentation of the next item. Yes/no judgements were recorded on the computer after each recognition item. For ease of analysis, strings that obeyed the artificial grammar are referred to as 'old' items and non-grammatical strings are referred to as 'new' items.

4.1.1.4. Pupil Recording

Maximum pupil-size was recorded from the right eye during each recognition period. A Pupil Dilation Ratio (PDR; see Chapter 2, section 2.1.2.1) was calculated expressing the maximum pupil-size for each 2000ms recognition trial as a proportion of the maximum pupil-size during that trial's 250ms baseline.

4.1.2. Results

4.1.2.1. Behavioural Data

The proportion of correct responses to old and new items was calculated for each condition. In the explicit condition, average hit rate (correctly identified old items) was 71.6%, average correct rejection rate (correctly identified new items) was 71.6%, false alarm rate (incorrectly identified new items) was 28.4% and miss rate (incorrectly identified old items) was 28.4%. Due to a programming error, participants' responses in the implicit condition were not recorded and the percentages could not be calculated.

4.1.2.2. Pupil-Size Data

Average PDR for old and new items was calculated for both conditions. As PDR is a function of baseline pupil-size, baseline pupil-sizes for old and new items in each condition were compared to ensure that differences in PDR were not due to baseline differences. The difference was not significant ($F(1,23) = 0.801, p > .05, \eta_p^2 = .034$).

A 2 x 2 repeated measures ANOVA of PDR with within-subject factors of item-type (old vs. new) and condition (implicit vs. explicit) revealed no significant main effect of item-type ($F(1,23) = 1.236$, $MSE < .001$, $p > .05$, $\eta_p^2 = .05$), but a significant main effect of condition ($F(1,23) = 68.00$, $MSE = .001$, $p < .001$, $\eta_p^2 = .75$) – across both old and new items participants' pupils dilated to a greater extent in the explicit condition ($M = 1.132$, $SD = 0.046$) than in the implicit condition ($M = 1.071$, $SD = 0.028$, $t(23) = 8.25$, $p < .001$, $r = .747$). The interaction between condition and item-type failed to reach significance ($F(1,23) = 2.184$, $MSE < 0.001$, $p = .153$, $\eta_p^2 = .087$), however, as is clear from Figure 4-2, there appears to be a difference between PDR for old and new items in the explicit but not implicit condition. *A priori* t-tests show that in line with previous findings PDR to old items ($M = 1.137$, $SD = 0.048$) in the explicit condition is larger than PDR to new items ($M = 1.127$, $SD = 0.048$, $t(23) = 2.024$, $p < 0.05$, $r = .151$; Figure 4-2).

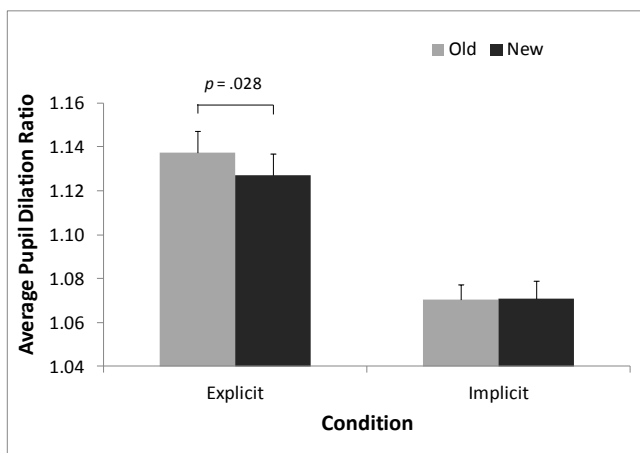


Figure 4-2: Pupil dilation ratio for old and new stimuli in both conditions. Error bars show standard error of mean.

4.1.3. Discussion

Experiment 3 replicated the PONE in the explicit recognition condition – participants' pupil-sizes increased to a greater extent when they viewed old items compared to

new items. The other key finding from this experiment is that there was no PONE in the implicit condition – average pupil dilations for grammatical and ungrammatical strings were the same. This suggests either that the PONE may be associated with recollection but not familiarity, or that insufficient levels of familiarity were generated by the grammatical strings to produce a measurable increase in PDR compared to non-grammatical strings.

However, the absence of behavioural data for the implicit condition makes further interpretation of the data difficult as it is not possible to restrict the analysis to correctly identified old and new items, which are presumed to be the best attended to stimuli and most likely to be engaging recognition memory processes. It is also not possible to determine whether the experimental design was effective and implicit learning took place, because we don't know whether "old" grammars were correctly judged to be grammatical at a level significantly above chance. Therefore the experiment was re-run with 23 participants, using the same stimuli and design.

4.2. Experiment 4 – Implicit Grammar Replication

4.2.1. Method

4.2.1.1. Participants

Twenty-three participants (3 male; age range: 18.9-49.1, $M = 21.9$, $SD = 6.06$) with normal or corrected-to-normal vision were recruited from the psychology course-credit and subject pools at the University of Sussex, and through personal contact.

4.2.1.2. Materials/Apparatus/Design/Procedure

As detailed for Experiment 3.

4.2.2. Results

4.2.2.1. Behavioural Data

The proportion of correct responses to old and new items was calculated for each condition. In the explicit condition, average hit rate (correctly identified old items) was 74.6%, and average correct rejection rate (correctly identified new items) was 72.8%. Examination of implicit data reveals that participants performed significantly above chance when correctly judging old strings as old (61.9%, $t(22) = 3.24$, $p < .01$, $r = .32$) and new strings as new (64.3%, $t(22) = 4.10$, $p < .001$, $r = .43$), suggesting that participants were able to learn the artificial grammar to some extent (see Table 4-1).

	Explicit		Implicit	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Hits	.746	.115	.619	.176
Correct Rejections	.728	.136	.643	.168
False Alarms	.272	.136	.355	.167
Misses	.254	.115	.381	.176

Table 4-1: Proportion of stimuli judgments for both conditions.

The proportion of correctly identified items was calculated for each condition. A 2 x 2 repeated-measures ANOVA with within-subject factors of item-type (old vs. new) and condition (implicit vs. explicit) showed a significant main effect of condition ($F(1,22) = 8.54$, $MSE = 0.03$, $p < .01$, $\eta_p^2 = .28$) – on average, participants identified more items correctly in the explicit than in the implicit condition – but no significant main effect of item-type ($F(1,22) = 0.011$, $MSE = 0.016$, $p > .05$, $\eta_p^2 = .001$) and no significant interaction ($F(1,22) = 1.24$, $MSE = 0.009$, $p > .05$, $\eta_p^2 = .054$; see Figure 4-3).

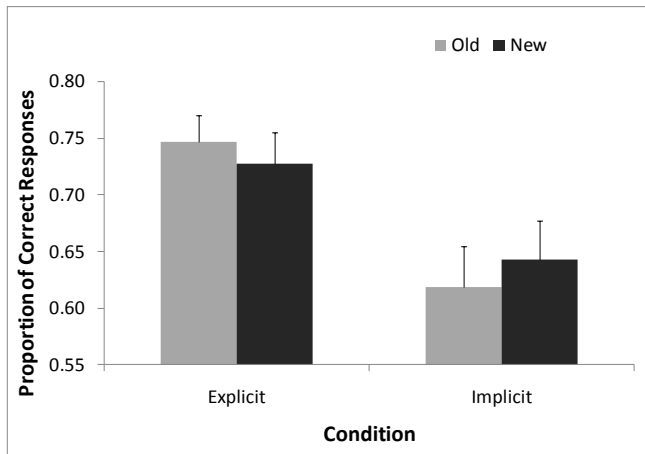


Figure 4-3: Proportion of correct responses. Error bars show standard error of mean.

4.2.2.2. Pupil-Size Data

Average PDR for old and new items was calculated for each condition. As PDR is a function of baseline pupil-size, baseline pupil-sizes for old and new items in each condition were compared to ensure that differences in PDR were not due to baseline differences. The difference was not significant ($F(1,22) = 0.943$, $p > .05$, $\eta_p^2 = .041$).

A 2 x 2 repeated-measures ANOVA of PDR for all trials with within-subject factors of item-type (old vs. new) and condition (implicit vs. explicit) revealed a significant main effect of item-type ($F(1,22) = 5.34$, $MSE < .001$, $p < .05$, $\eta_p^2 = .195$) – participants' pupils dilated to a greater extent to old items ($M = 1.066$, $SD = 0.034$) than to new items ($M = 1.061$, $SD = 0.033$, $t(22) = 2.31$, $p < .05$, $r = .195$). The main effect of condition was also significant ($F(1,22) = 28.2$, $MSE = .001$, $p < .001$, $\eta_p^2 = .562$) – participants' pupils dilated to a greater extent in the explicit condition ($M = 1.083$, $SD = 0.046$) than in the implicit condition ($M = 1.044$, $SD = 0.026$, $t(22) = 5.31$, $p < .001$, $r = .562$).

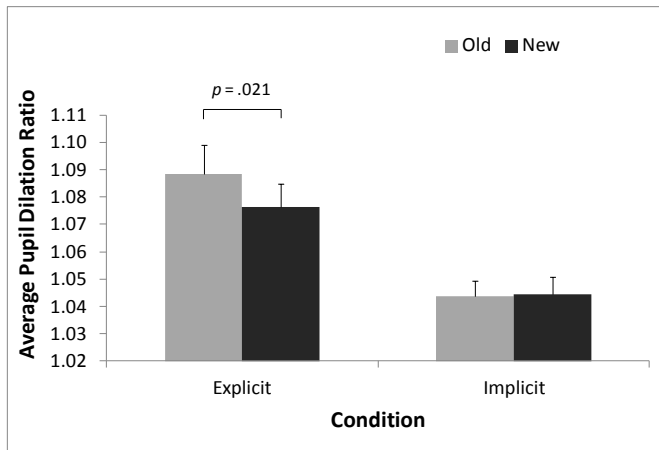


Figure 4-4: Pupil dilation ratio to old and new items in each condition. Error bars show standard error of mean.

The interaction between condition and item-type just failed to reach significance ($F(1,22) = 3.13$, $MSE < 0.001$, $p = .09$, $\eta_p^2 = .125$). However, t-tests show that in the explicit condition PDR to old items ($M = 1.089$, $SD = 0.048$) is larger than PDR to new items ($M = 1.076$, $SD = 0.046$, $t(22) = 2.49$, $p < .05$, $r = .220$) whereas this difference is not significant in the implicit condition (old: $M = 1.044$, $SD = 0.027$; new: $M = 1.045$, $SD = 0.029$; $t(22) = -0.193$, $p > .05$, $r = .002$; see Figure 4-4).

The analysis was repeated, but restricted to items that were correctly identified. The main effect of condition remained significant ($F(1,22) = 25.2$, $MSE = 0.001$, $p < .001$, $\eta_p^2 = .534$) – average PDR was larger in the explicit condition than in the implicit condition. The main effect of item-type also remained significant ($F(1,22) = 5.04$, $MSE < 0.001$, $p < .05$, $\eta_p^2 = .186$) – participants' pupils dilated to a greater extent to old stimuli than for new stimuli.

Importantly the item-type by condition interaction reached significance ($F(1,22) = 9.03$, $MSE < 0.001$, $p < .01$, $\eta_p^2 = .291$) – the difference in PDR between correctly identified old and new items in the explicit condition ($M = 0.022$, $SD = 0.033$) was larger than the difference in PDR between correctly identified old and new items in

the implicit condition ($M = -0.003$, $SD = 0.022$, $t(22) = 3.00$, $p < .01$, $r = .290$) – the PONE was present in the explicit but not the implicit condition (Figure 4-5).

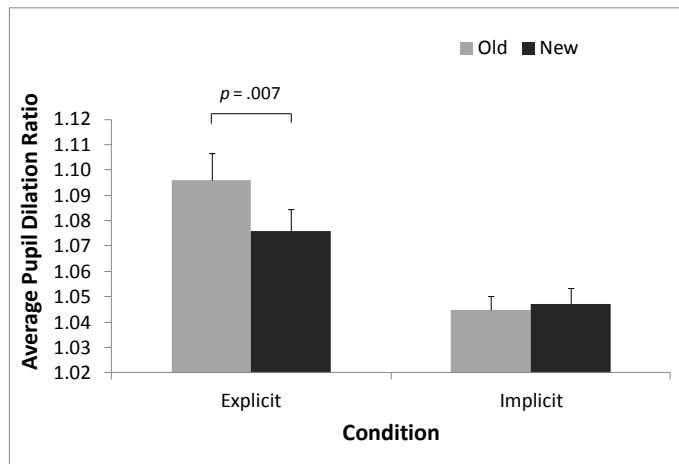


Figure 4-5: Pupil dilation ratio to correctly identified old and new items in each condition. Error bars show standard error of mean.

4.3. General Discussion

Experiments 3 and 4 sought to determine whether the relative increase in pupil-size that occurs when participants view previously learned items during a recognition test, compared to novel items, also occurs when participants make “old/new” decisions that reflect implicit learning as opposed to explicit recognition of previously encountered items. The key finding is that whilst the PONE was again observed in the explicit recognition conditions in both experiments, there was no evidence for any increase in pupil-size when participants were exposed to grammatical compared to ungrammatical letter strings.

Behavioural measures showed that in both conditions participants were attending to and learning stimuli reasonably well. Across the explicit conditions of Experiments 3 and 4 participants were able to correctly identify 72.8% of old words, and to correctly judge 72.4% of new words, consistent with the literature (e.g., Yonelinas, 1994).

Behavioural measures in Experiment 4 showed that participants were attending to the

stimuli and had been able to learn the artificial grammar to some extent, they performed above chance when correctly judging 61.9% of old strings and 64.3% of new strings, consistent with the degree of learning shown using these stimuli in recent studies (cf., Scott, & Dienes, 2008; 2009).

The PONE was not present in the implicit condition of either experiment – average pupil dilations for grammatical and ungrammatical strings were the same. This was the case even when the analysis was restricted in Experiment 4 to those items that participants successfully identified as grammatical. Artificial Grammar Learning (AGL) was used in order to investigate the PONE in a recognition situation in which familiarity, but not recollective processes were involved. Scott and Dienes (2008) showed that participants make “old” decisions in implicit grammar learning tasks based on a sense familiarity, and that stronger perceived familiarity ratings lead to more correct hits. They suggest this may reflect unconscious recognition of the stimuli’s features as a result of implicit learning during the learning phase. The strings presented in the recognition phase of AGL tasks were all new, therefore no recall of individual whole strings should occur (although the possibility that small fragments are explicitly recognised remains).

According to the signal-detection unequal-variance theory, both familiarity and recollection can vary in strength and the greater their combined strength, the greater the memory signal. Recognition results only if this signal exceeds a recognition threshold (Wixted, 2007a; Wixted, & Stretch, 2004; Otero et al., 2011). If, as Otero et al. (2011) have argued, the PONE reflects an aggregate of both recollection and familiarity signals, the PDR should be greater for grammatical strings than non-grammatical strings, due to a greater sense of familiarity in the absence of conscious recognition (Scott, & Dienes, 2008) for these items. The absence of the PONE in the

implicit condition suggests that it may reflect recollective but not familiarity processes. This suggests that genuine amnesiacs may not show the PONE in the absence of explicit memory. Alternatively, the PONE might reflect both recollection and familiarity signals, but in the absence of recollection in the implicit condition the level of familiarity achieved was insufficient to produce a measureable effect on pupil-size.

As discussed in Chapter 3, pupil response has already been found to indicate intact implicit memory in amnesic patients. Our results suggest that the increase in pupil-size that occurs when participants encounter correctly recognised old items reflects neurocognitive processes associated with explicit, but not implicit, recognition memory. This could indicate that the pupillary response is a function of task demands (recognition memory test) as opposed to being the automatic consequence of being exposed to items previously encountered during a learning phase (confirming the findings of Experiments 1 and 2). The larger PDR which occurs to new items in the explicit condition compared to 'new' (ungrammatical) items in the implicit condition may occur because the grammatical decision in the implicit condition does not involve recognition memory processes, whereas the correct rejection of a new item in the explicit condition may involve an effortful memory search, leading to effort-related increases in pupil-size (e.g. Kahneman, 1973), and studies have found that rejection typically takes longer than correct recognition (e.g., Ratcliff, & Murdock, 1976).

In Chapter 5, the issue of whether the PONE reflects automatic or voluntary mnemonic processes is explored further in a series of experiments that seek to establish the extent to which it is associated with item-type (old vs. new) or participant response (old vs. new).

5. Malingering and the Old/New Response

Is the PONE Under Voluntary Control?

Building on the findings of Experiments 1 and 2, the results of Experiments 3 and 4 suggest that the increase in pupil-size that occurs when participants encounter previously studied items, and recognise them as old, reflects neurocognitive processes associated with explicit, but not implicit recognition memory. This could indicate that the PONE represents processes involved in making an explicit recognition memory judgement, for example a memory strength signal, which is larger for old items than new items, and suggests that conscious awareness is an important factor in the PONE. Therefore, in Chapter 5 this idea is extended to explore whether the PONE reflects voluntary mnemonic processes, in three experiments that seek to establish the extent to which it is associated with item-type (old vs. new) or participant response (old vs. new). Experiments 5 and 6 follow the logic of studies that have attempted to exploit the ERP old/new effect as an index of feigned amnesia (malingering), to investigate the effect of asking participants to deliberately give wrong answers during the recognition phase. Experiment 7 attempts to genuinely impair recognition memory performance in healthy participants by dividing attention.

5.1. Malingering and Deception

5.1.1. Background

The nature of deception and how to detect it has long concerned scientists, philosophers, clinicians and forensic professionals. There is also a significant commercial interest in creating a simple, reliable, and accurate evidence-based method of lie-detection that could be used to identify individuals attempting to deceive

or malingering for personal or financial gain, or to evade criminal prosecution. The ultimate goal of lie-detection is a method that doesn't rely on subjective assessment, and is not subject to the same interpersonal manipulations as people.

A particularly complex type of lying is malingering, the deliberate fabrication or exaggeration of physical or psychological symptoms for secondary gain, such as financial compensation, evading a penalty or conviction, gaining drugs, or avoiding military service (Hutchinson, 2001; Tugcu, 2010). Adults, adolescents and children may feign a wide range of conditions including movement disorders, sensory issues, epilepsy, loss of consciousness, and neurological deficits (Kasikci, & Bek, 2010). It has been estimated that the incidence of malingering is approximately 12% in patient populations, rising to 40% in patients seeking financial compensation (see Hutchinson, 2001; Larrabee, 2003; Mittenberg, Patton, Conyock, & Condit, 2002). There may be elaborate motivations behind malingering and, as Hutchinson (2001) observes, the boundaries between exaggeration and fabrication, and deliberate and unintentional malingering are not clear. For this reason, some believe that within a biopsychosocial model, use of a value-laden term like "malingering" is prejudicial, preventing best practice, and that patients with such "disorders of simulation" have a genuine need for help (Hutchinson, 2001; Ucar, & Atac, 2010).

Malingering does not have well-established diagnostic criteria, primarily due to clinicians' reliance on patients' self-reported symptoms (Hutchinson, 2001). This is particularly difficult in cognitive neuropsychology, where there may be no objectively measurable physical symptoms, and patients may fabricate language impairments (e.g., Cottingham, & Boone, 2010), posttraumatic stress disorder (e.g., Merten, Thies, Schneider, & Stevens, 2009), low intelligence, psychosis, amnesia, or affective disorders (e.g., Slick, Sherman, & Iverson, 1999). Most conditions are assessed

through clinical interview and/or the administration of neuropsychological tests designed to measure the degree of difficulty or impairment. Feigned poor performance is difficult to detect, and presents a real challenge to practicing clinicians (see Hutchinson, 2001). According to Binder and Rohling (1996), patients seeking financial compensation following closed head injury perform more poorly on neuropsychological assessments than do patients who are not seeking financial reward by around one half of a standard deviation. It has been shown that effort and cooperation, rather than brain injury, accounts for up to 50% of the variance in results (Rohling, Green, Allen, & Lees-Haley, 2000). Mittenberg et al. (2002) have argued that survey-based studies which attempt to identify the base-rate of malingering are likely to underestimate the true rate, as the ability of clinicians to detect malingerers is not 100%.

In response to these issues, psychologists have developed instruments which index the likelihood of malingering, for example the Fake Bad Scale (FBS; Lees-Haley, English, & Glenn, 1991) which relies on malingerers' desire to appear healthy and honest, and is sensitive to "illogical symptom histories" (Greiffenstein, Baker, Gola, Donders, & Miller, 2002) or insufficient cognitive effort in litigants with mild traumatic head-injury (Ross, Millis, Krukowski, Putnam, & Adams, 2004; Slick et al., 1996).

5.1.1.1. Malingered Memory-Impairment

A commonly malingered cognitive deficit is memory-impairment (Rüsseler, Brett, Klaue, Sailer, & Münte, 2008; Samuel, & Mittenberg, 2005). Memory is a key cognitive function, and disruption to the memory systems impacts massively on a person's quality of life, which has implications for the types of service they receive and, in the case of accidents, the amount of compensation. It is widely known to the general public that memory-impairment is a frequent outcome of a head injury such

as concussion (Gouvier, Prestholdt, & Warner, 1988), and even patients with genuine memory-impairment may exaggerate their symptoms when seeking compensation (Yochim, Kane, Horning, & Pepin, 2010).

Therefore existing memory measures have also been adapted to measure the likelihood of malingering, such as the expanded version of the Auditory Verbal Learning Test (Barrash, Suhr, & Manzel, 2004) and the Wechsler Memory Scale Third Edition (WMS-III; Wechsler, 1997). In addition specific malingering tests have been developed, such as the Test of Memory Malingering (TOMM; Tombaugh, 1996), Word Memory Test (Green, & Astner, 1995) and the Victoria Symptom Validity Test (VSVT; Slick, Hopp, Strauss, & Thompson, 1997). These rely on patterns of performance that are inconsistent with the performance of genuinely memory-impaired individuals, for example poorer recognition performance than delayed recall, and absence of memory effects such as the primacy effect (Barrash et al., 2004).

The WMS-III has a Rarely Missed Index in its scoring system, of items that are unlikely to be forgotten by participants with genuine memory-impairment, classifying the genuinely neurocognitively-impaired from those fabricating memory difficulties with an accuracy of 98% (Wechsler, 1997), and head injured patients from those with substance abuse with an accuracy of 95% (Miller, Ryan, Carruthers, & Cluff, 2004). However, subsequent studies including genuine patients, participants exaggerating symptoms, and head-injury litigants, found less desirable overall classification accuracies of 75-89% (Lange, Senior, Douglas, & Dawes, 2003; Langeluddecke, 2004), specificities of 87-91% (correctly identifying non-malingers as genuine; Lange et al., 2003; Lange, Sullivan, & Anderson, 2005; Swihart, Harris, & Hatcher, 2008), and much lower sensitivities of only 18-25% (correctly identifying a malingerer as such; Lange et al., 2005; Swihart et al., 2008). Axelrod, Barlow and Paradee (2010)

indicate that in the case of memory-impairment severe enough to cause random responses, 69% of respondents would fall within the range of suboptimal effort and may be falsely accused of malingering.

The task of assessing memory-impairment is even more of a challenge among forensic populations (Denney, 2007), because when neuropsychological measures, which have been standardised on a civil population, are used to assess criminal populations, they show a negative bias toward a more impaired performance than would be predicted by their general cognitive ability (Ardolf, Denney, & Houston, 2007; Franzen, & Iverson, 2000). It is estimated that 70% of forensic patients are thought to adapt their presentation when assessed by a clinical neuropsychologist (Heilbrun, Bennett, White, & Kelly, 1990), with 25-50% of murder and manslaughter suspects claiming crime-related amnesia (Pujol, & Kopelman, 2003). Although some genuine memory losses are reported for the period in which the crime was alleged, due to intoxication, seizure, or sleep disorder (Bourget, & Whitehurst, 2007), in a study of over 300 convicted prisoners, Cima, Nijman, Merckelbach, Kremer and Hollnack (2004) found none whom they considered to have genuine trauma dissociation amnesia, stating that such claims should be approached with caution.

Clearly, the severe implications of being wrongly convicted, or acquitted, of a crime require accurate methods of detecting crime-related amnesia (Bourget, & Whitehurst, 2007; Merckelbach, & Christianson, 2007). A concerning example of the flaws in neuropsychological tests comes from Galappathie and Vakili (2009), who report a case study of a man attempting to evade a conviction for murder by feigning memory-impairment. Mr X initially “passed” the TOMM, but six months’ observation showed his abilities to learn new information and recall past information to be functioning well above the level indicated by neuropsychological tests. He later “failed” a second

administration of the TOMM, was deemed to be making insufficient effort and his trial was resumed. O'Bryant and Lucas (2006) showed that although the sensitivity of the TOMM is 98%, its specificity is only 78%, meaning that only 3 out of 4 people who "pass" the TOMM are genuinely making enough effort.

Unfortunately neuropsychological tests may also be vulnerable to countermeasures (attempts by the individual to distort the results in their favour), such as coaching, where participants are given detailed information about the symptoms of memory-impairment or strategies to avoid detection (e.g., Bauer, & McCaffrey, 2006; Jelicic, Merckelbach, Candel, & Geraerts, 2007). Although layperson ideas about amnesia may be inaccurately based on examples from films (Baxendale, 2004), as tests advance, so too does the sophistication of malingering strategies. In an arms race with test developers, malingerers can use the internet to access up-to-date information such as pass/suspicious/fail cut-off scores, details of the scoring system, and example stimuli (Bauer, & McCaffrey, 2006). Differences in performance levels, reaction times and responses to feedback between genuinely memory-impaired, non-memory-impaired and malingering populations on a range of different tests can also be researched. In this way potential malingerers can determine situations and tests on which memory-impaired patients would actually be expected to perform well, such as on the Amsterdam Short Term Memory test (ASTM; Schagen, Schmand, de Sterke, & Lindeboom, 1997; Bauer, & McCaffrey, 2006; Jelicic, Merckelbach, Candel, & Geraerts, 2007). Bauer and McCaffrey (2006) found that the TOMM was particularly well documented on the internet and suggest that its simpler format may be more conducive to malingering than tests that measure multiple domains or include reverse-scored items of bizarre or atypical symptoms not usually associated with genuine impairment, such as in the Structured Inventory of the Malingered Symptomatology (SIMS; Smith, & Burger, 1997).

In short, most tests of malingering based on participants' responses rely on malingerers making too many errors and taking too long to respond compared to genuinely memory-impaired individuals. As such, careful research about symptoms and how the test works could result in an individual feigning believable poor performance. Although the administration of several tests increases the likelihood of detection, this is time consuming. Peter, Merten, Merckelbach and Oswald (2010) found that after warning malingering simulators not to exaggerate symptoms, despite the improved sensitivity and specificity of using multiple validated instruments, three out of twenty malingerers were able to evade detection.

5.1.1.2. Psychophysiological Detection of Deception

Memory is subjective, and behavioural responses rely on the ability and desire to respond accurately. An obvious literature for neuropsychological professionals to turn to when seeking solutions to problems of memory malingering is that of the psychophysiology of deception detection. Psychophysiological techniques may offer a more suitable way of assessing memory performance without reliance on self-report (see Chapter 1, section 1.3.2).

The attempt to identify measurable psychophysiological markers of deception has a long history. As long ago as 1000 BC physiological responses to stress were documented as being used to identify liars (for a historical overview of the nature of deception and lie-detection, see Ford, 2006). The first four-channel polygraph device was developed in 1932 to assess heart rate, blood pressure, respiratory rate and GSR (Saxe, Dougherty, & Cross, 1985). Typically this approach operates on the basis that, for the majority of people, deception is generally more cognitively demanding, and more anxiety provoking, than telling the truth. This results in an involuntary difference in the psychophysiological response to incriminating stimuli

compared to non-incriminating stimuli in a guilty person. The polygraph works on the assumption that changes in these autonomic responses indicate dishonesty. It does not take into account other influences, such as emotional upset, fatigue or medication, and is open to manipulation by individuals who understand the process and use countermeasures; for example, thinking anxiety-provoking thoughts whilst their baseline levels are being measured (Ford, 2006). Individuals with antisocial personality disorder or psychopathy may be more able to pass polygraph tests whilst lying (Verschuere, Crombeza, Kostera, & De Clercq, 2007). This is proposed to be due to autonomic hypo-responsivity in psychopathic individuals, in particular a reduced skin conductance, and a reduced resting heart rate (Verschuere et al., 2007).

Contemporary polygraph lie detection, which could be seen to represent the state of the art, still measures autonomic nervous system activity via breathing rate and pattern, GSR, blood pressure, and heart rate, but its validity and reliability are contested (Lykken, 1998). Reviews suggest that accuracy varies between 50% and 90% (Brett, Phillips, & Beary, 1986; Stern, 2003) with a converging approximation of around 75% sensitivity (correctly identifying guilty persons) and 65% specificity (correctly identifying innocent persons) (Brett et al., 1986). Polygraph evidence is usually not permitted in a court of law (Ford, 2006). A recent report by the United States National Research Council (NRC) concluded that there was insufficient evidence that the polygraph is able to accurately detect deception, and that, "Countermeasures pose a serious threat to the performance of polygraph testing because all the physiological indicators measured by the polygraph can be altered by conscious efforts through cognitive or physical means" (National Research Council, 2003, p. 4), highlighting the need for alternative methods. For these reasons, modern interest in deception-detection has moved on from measures of peripheral nervous

system activity, such as polygraphs, to more direct assessment of the central nervous system, using techniques such as fMRI or EEG.

Crucial to detecting deception through physiological changes, is understanding of the neurological underpinnings of deliberate deception. The immense progress that has been made in psychophysiological techniques in the last few decades has begun to untangle these processes and direct future research. Psychophysiological techniques measure the physiological responses generated by cognitive activities, using markers such as blood flow in the brain, electrical or magnetic field changes, and pupil dilation. Multiple techniques have been developed to assess subtle task-induced changes in psychophysiological markers in the laboratory. These approaches share goals, such as gaining understanding of 'hidden' psychological processes, but they differ in aspects such as methodology, equipment, resolution, which structures, signals and processes can be measured, how invasive participation may be, and the types of participants who are suitable for analysis. Different techniques bring complementarities to the topic through their strengths and weaknesses, and different populations, such as healthy adult, brain injured, elderly, and developmental, contribute different insights.

The majority of techniques used to assess central nervous activity can be broadly divided into real time recording techniques that directly tap the activity of the neurons, such as ERPs and MEG, and indirect techniques that monitor cognitive activity by the haemodynamic and metabolic changes within the brain, such as fMRI and Positron Emission Tomography (PET). Diffuse Optical Tomography (DOT) techniques such as Near Infrared Spectroscopy (NIRS) and Event-Related Optical Signal (EROS) are non-invasive, cheap and relatively portable, and may have the potential to measure

directly and indirectly (Irani, Platek, Bunce, Ruocco, & Chute, 2007), but have not yet been extensively applied to the topic of deception and won't be considered here.

In conducting empirical tests of the efficacy of psychophysiological indices of deception, it is important to consider the experimental protocols involved. Studies in the detection of deception have used various experimental deception protocols. One common paradigm involves asking a participant to commit or observe a small crime in another room, and then conceal their involvement in the incident from the experimenter (e.g., Berrien, & Huntington, 1943; Mohamed et al., 2006). This Guilty Knowledge Test (GKT), or Concealed Information Test (CIT), has been used extensively with polygraph tests and in applied forensic settings (Ben-Shakur, & Elaad, 2003). It involves forced-choice answers to questions about the crime under investigation, where one "relevant" answer contains information about the incident, and several "neutral" answers act as control items. Neutral answers are chosen so that they are indistinguishable from the relevant answer to innocent individuals, for example: "You stole money from my wallet. Was it £20, £100, £10, £50, £30, £70?" (Engelhard, Merckelbach, & van den Hout, 2003). Participants are asked to respond "no" to each item, and guilt is inferred if individuals consistently display larger psychophysiological responses (for example GSR or ERPs) when giving the relevant answer compared to neutral answers (Ben-Shakur, & Elaad, 2003). A variation of the GKT uses three different item-types in a task similar to a standard recognition task: 'target' items, provided for the participant to learn prior to the test (e.g., farm animals) and which require a "yes" response, and two types of non-learned items that require a "no" response - 'probe' items, representing knowledge relevant to the investigation that only the guilty person would know (e.g., "ring" when the mock crime was stealing a ring), and 'irrelevants', neutral items not relevant to the investigation (e.g., 5 other items of jewellery) (Ford, 2006). An innocent person will respond to

probes and irrelevant items in the same way as probes are no more meaningful than the irrelevant items, however guilt is inferred if psychophysiological responses to probes are similar to targets, which are meaningful (Ford, 2006). Reviews of the GKT have shown it to have a sensitivity of 70%-85% (Ben-Shakhur, & Elaad, 2003) and specificity of 80%-82% (MacLaren, 2001).

As the NRC point out, measures such as breathing rate and GSR can be manipulated consciously. In an attempt to find measures that are not so susceptible to countermeasures, researchers have used ERPs to try to detect deception. Abootalebi, Moradi and Khalilzadeh (2006) carried out a single-probe GKT whilst recording ERPs. In two tasks participants chose whether they were 'guilty' or 'innocent'. Prior to being questioned, guilty participants looked in a box containing an object whilst the examiner was outside the room. The test contained five pictures of objects, one a target object that had been shown to all participants, one a probe (the object in the box) and three objects that had not been seen by any of the participants. Participants then had to indicate whether they had seen the object in the picture and guilty participants had to hide their knowledge of the probe by responding "no". Abootalebi et al. (2006) measured the P300, an ERP thought to represent higher cognitive responses, to cognitively salient, distinct, learned or unexpected stimuli (see Polich, 2007, for a review; Molnar, 1994; Sutton, Braren, Zubin, & John, 1965). Due to the fact that each set of stimuli were presented 30 times, giving a single average P300 ERP per item-type per participant, the authors used a bootstrapping (Wasserman, & Bockenholt, 1989) algorithm to estimate the distribution of average P300 waves and create a "guilt threshold" with which to assess an individual's innocence. Using bootstrapped amplitude differences to classify participants they were able to correctly detect 74% of guilty participants, but noted that a guilty

individual with a poor P300 response to probe stimuli would remain undetected and be classified as innocent.

Whilst the psychophysiology of deception literature is primarily concerned with forensic rather than neuropsychological applications, some studies have looked directly at malingering using psychophysiological methods. For example, Wu, Allen, Goodrich-Hunsaker, Hopkins and Bigler (2010) showed with fMRI scans that, in spite of considerable brain damage, the same brain regions were activated during the WMT as in controls, whereas different patterns were seen for participants who were asked to simulate incomplete effort, suggesting that fMRI could be used to differentiate malingered and genuine performance.

ERPs have also been used to investigate feigned memory-impairment. Rosenfeld, Ellwanger and Sweet (1995) used an oddball paradigm in which autobiographical items such as participants' dates of birth, phone numbers and mothers' maiden names were presented on a screen among eight unrelated non-oddball stimuli of the same type. Participants were required to repeat the item out loud to ensure that "malingered" was not achieved simply by participants avoiding looking at the items. They found an enhanced P300 in response to these uncommon, personally relevant items, regardless of overt recognition response. Participants were classed as guilty if their oddball to average non-oddball amplitude ratio was larger than 1.5, and no individual non-oddball amplitude was more than 20% larger than the average non-oddball amplitude. Using this method the authors were able to identify 92% of malingerers for birthdates and phone-numbers and 77% for mothers' maiden names.

Whilst ERP based techniques show some promise, they are costly in terms of both equipment and time. In addition, as with standard polygraph techniques, countermeasures to defeat P300 as an index of deception are easily learned, such as

moving fingers or toes, or visualising being slapped, whilst responding to irrelevant stimuli, to generate a P300 similar to those evoked by relevant and probe stimuli (Rosenfeld, Soskins, Bosh, & Ryan, 2004). Rosenfeld et al. (2008) report development of a new, countermeasure-resistant P300-based method for detecting concealed information, The Complex Trial Protocol, which identifies attempted countermeasure use from extended reaction times (Winograd, & Rosenfeld, 2011).

Lykken (1959; 1960) devised the GKT to test recognition memory, rather than deception-related stress, and the neuroelectric changes associated with mnemonic processes have been extensively documented in the ERP literature. For example a larger frontal negative going component (FN400) and larger parietal late positive component (LPC) have been found in response to the presentation of old learned items compared to new items during a recognition memory test (Warren, 1980; van Hooff et al., 1996; Friedman, & Johnson, 2000; see Chapter 6, section 6.1.1 for more detailed discussion of these components). This ERP old/new effect provides researchers with a more direct strategy with which to attempt to identify people who feign memory loss (e.g., Browndyke et al., 2008; Tardif, Barry, Fox, & Johnstone, 2000; van Hooff, Sargeant, Foster, & Schmand, 2009).

For example, Tardif et al. (2000) reasoned that if the ERP old/new effect is not under conscious control then it should be detectable in people feigning amnesia when they claim that previously encountered old information is actually “new”. In their experiment, participants learned a set of words before completing a recognition memory test. Half the participants were given standard test instructions to respond “old” to the learned words and “new” to new words. The other half were asked to perform deliberately poorly on the test. Although the behavioural performance was worse for the malingering group, the ERPs revealed an old/new effect comparable in

magnitude and topography to the control group, suggesting that the malingering group did in fact have intact recognition of the learned words. In addition, the malingering group did not show differences in response latency to correct and incorrect responses, and demonstrated an early (190-320ms) old/new P2 difference that was not present in the truthful group – which may represent processing differences involved in the act of malingering (Tardif et al., 2000). Classification was carried out using a direct discriminant function analysis of differences in reaction time, LPC old/new effect and P2 old/new effect, and correctly identified 84.2% of malingerers and 78.9% of the control group. However, Allen and Mertens (2009) suggest that memory distortion limits the accuracy of this type of approach. Their study showed that under certain conditions, patterns of neural activity for true and genuine false recognition are indistinguishable.

5.1.1.3. Pupillometry Studies

As discussed in Chapter 1, section 1.2.3.7, despite being under autonomic control, relatively little research has been conducted into the activity of the pupil during lying (Janisse, 1977). However, the eyes in general have long been viewed as non-verbal indicators of truth-telling. In their meta-analysis of 158 different cues to deception in 120 independent samples spanning 60 years of research, DePaulo et al. (2003) found several aspects of the eyes had been associated with lying, such as the amount of eye contact (avoidance), gaze direction, blink rate (linked to arousal and cognition, see Chapter 1, section 1.2.4.2) and pupil dilation (also linked to arousal and cognitive effort, see Chapter 1, sections 1.2.2 and 1.2.2.1). Of these, only pupil-size demonstrated a statistically significant effect size ($d = .39$, $p < .05$; DePaulo et al., 2003), yet the relationship between pupil-size and deception has received comparatively little attention, possibly because such subtle changes are much harder

to detect and measure accurately or easily in real world situations, and therefore may be less consciously linked with dishonesty.

Elaad (2009) investigated male police officers', prisoners' and laypersons' beliefs regarding their own and others' nonverbal abilities at detecting and telling lies. The questionnaire included perceived cues (e.g., "Liars are more gaze aversive") and actual cues to deception (e.g., "Lies are accompanied by pupil dilation") and found that prisoners (and in particular older prisoners) were more aware of pupil dilation as a cue to deception (50%) than lay persons (17%), or even police officers (21%). Around half the participants in all three groups were aware that, "Lies contain more negative statements". Elaad (2009) suggests that the increased awareness among prisoners may be due to increased exposure to and corrective feedback from lies.

In an early study Berrien and Huntington (1943) asked half of their participants to commit a small monetary theft as part of the experiment, and to later lie when questioned about it. They compared changes in their pupil-size with a group of participants with no knowledge of the crime, instructed to answer questions truthfully. Measurements were made using a telescope focussed on the eye and moved left and right by an observer as the pupil dilated or constricted; movements were transmitted to, and amplified by, a capillary pen on a polygraph. All participants were asked unrelated baseline questions, and questions relating to the crime; deceivers were told that they would be able to keep the money if their dishonesty remained undetected. Lying was associated with a greater incidence of slow dilations followed by quick constrictions and increased instability in pupil-size. They also found that pupil dilation appeared to be more specific to the liars than concurrent increases in blood pressure, and attributed this to the emotion evoked by being dishonest. However, in the

absence of any statistical analysis or objective quantification of the changes in pupil-size, this finding clearly requires replication with more modern techniques.

Later attempts using photographic methods confirmed increased pupil-size when participants lied about demographic information, were asked about “guilty information”, or viewed photographs of items from the scene of the “crime” (e.g., Heilveil, 1976; Janisse, 1973; Bradley, & Janisse, 1979; 1981; Janisse, & Bradley, 1980; Lubow, & Fein, 1996). Given the large body of evidence demonstrating that pupil-size increases with “arousal” and is correlated with other psychophysiological measures of arousal such as heart rate and GSR (see Chapter 1, section 1.2.6), researchers have often used the logic behind polygraph recording to investigate the potential of pupil-size as an index of deception in laboratory settings.

Lubow and Fein (1996) combined a GKT with pupil-size measurements to index processing load, arousal and emotional state. Participants were randomly allocated to either an innocent group or a group instructed to carry out a mock crime. Photographs of the crime-scene and aspects of the crime acted as probes, and participants were told that their pupil-size, eye movements and GSR would be measured during the ‘interrogation’. Guilty participants had larger pupil-sizes in response to probe items than to control items, whereas innocent participants’ pupil-sizes were comparable between probe and control items. Participants with guilty knowledge also had larger pupil-sizes to control items compared to innocent participants. The experimenters were able to correctly classify 50% guilty and 100% innocent participants using only differences in pupil-size, comparable to electrodermal data (55% and 93%).

Webb et al. (2009) examined pupil diameter using a mock crime experiment where half the participants stole money in the lab and all participants were later given a

Comparison Question Test (CQT). Like the GKT, the CQT asks questions relevant to the crime (e.g., "Did you take any of the missing money?"), but also probable-lie questions that are vague and difficult to answer when trying to appear honest (e.g., "Before the age of 30, did you ever take something that did not belong to you?"). The rationale is that innocent participants will react more to the probable-lie questions because they can honestly answer the relevant questions, whereas guilty participants will react more strongly to the relevant questions. The CQT is controversial, because results may reflect surprise, anxiety or stress as much as deception, requiring subjective interpretation by the investigator (see Ben-Shakhar, & Furedy, 1990; Honts et al., 2005; Iacono, & Lykken, 2005), however it is used internationally in actual forensic cases (Raskin, & Honts, 2002). Webb et al. (2009) found that innocent participants showed larger increases in pupil-size to probable-lie than to relevant questions, whereas guilty participants showed similar increases in pupil-size to both question types. Regression analyses revealed that pupil-size was a significant predictor for deception, improving the adjusted R^2 from .39 to .46, however this increase only approached significance. The authors concluded that pupil-size could be used in place of blood pressure measurement in traditional polygraphy, but did not increase detection rates very much in addition to existing measures.

Using a different approach, Dionisio et al. (2001) asked participants to answer episodic and semantic questions relating to general knowledge ("What are the colours of the American flag?") or specific vignettes ("What was the name of the person in the story?"). They were required to answer twice, once honestly and once deceptively. In 92% of participants they observed significantly larger pupil dilation for both types of question when participants were lying. The researchers proposed that greater cognitive processing was associated with creating convincing deceptive responses than with genuine recall, thus leading to larger pupil-sizes (Dionisio et al., 2001).

Wang (2010) argues that deception studies often lack realism because their nature dictates participants' responses, and real-world motivations to malingering, such as strong financial or emotional incentives, are largely absent (Bauer, & McCaffrey, 2006). Wang, Spezio, and Camerer (2010) attempted to address these drawbacks in a biased-transmission game eye-tracking study, where the sender communicates biased information and gets more points for greater exaggerations if they successfully mislead their receiver. They found that senders' pupils dilate when sending deceptive messages compared to sending accurate messages, and that dilation increases with the magnitude of the deception. They suggested that this was because figuring out how much to deceive another player is cognitively difficult (Wang et al., 2010).

However psychophysiological approaches that rely on indices of increased "cognitive effort" and "arousal" are still vulnerable to countermeasures such as counting backwards in sevens or thinking anxiety-provoking thoughts whilst baseline levels are being measured, and attempting to ignore relevant items (Rosenfeld et al., 2004). In addition, if a person is able to lie easily without increased stress, then these measures would not be appropriate – for example, individuals with antisocial personality disorder or psychopathy are thought to pass polygraph tests whilst lying, due to the hyporesponsivity of their autonomic nervous system (Verschuere, Crombeza, Kostera, & De Clercq, 2007). It remains to be seen whether pupil-size is also susceptible to countermeasures in relation to deception. Ekman, Poikola, Mäkräinen, Takala and Hämäläinen (2008a; Ekman, Poikola, & Mäkräinen, 2008b) have designed a computer game that responds to player pupil-size and have had modest success in training participants to manipulate their own pupil-size by holding their breath, hurting themselves, reflecting on an emotional event, performing mental arithmetic or changing their point of focus in a (slow) paced task.

As yet the pupillometry research has not yet been extended to look at the malingering of memory-impairment. Given the comparative ease with which pupil measurements can now be made, it is important to establish whether an approach similar to the ERP old/new malingering studies (e.g., Tardiff et al., 2000; van Hooff et al., 2009) might be feasible using the PONE. Some early evidence suggests that, like the ERP old/new effect, the PONE is not under voluntary control during a memory task. Clark and Johnson (1970) informed participants that their pupil would increase or decrease in size during a short term memory task, or, in a control condition, did not mention pupil-size at all. They found that pupil-size increased to a similar extent in each condition, not only when participants had been told it would increase. If the PONE is not under voluntary control, and represents an automatic consequence of successful recognition, this could provide a method of detecting deception that is independent of emotional stress levels and cognitive effort.

This chapter presents a series of three experiments which artificially reduce recognition memory performance to explore the role of conscious awareness in the PONE, and the extent to which it may relate to conscious control. Experiment 5 looks at whether the PONE is under voluntary control by asking participants to simulate memory malingering to determine whether the pupil responses are aligned with item status (old or new) or participant response (old or new). In an attempt to further establish the effect of particular strategies on the PONE, Experiment 6 looks at different types of malingering strategy, including asking participants to provide incomplete effort, or to think their answer without indicating behaviourally to the experimenter. Finally Experiment 7 looks at whether dividing attention at learning and/or recognition reduces recognition performance (simulating memory-impairment) without participants using a malingering strategy, and how genuinely reduced memory performance affects the PONE.

5.2. Experiment 5 – Malingering and the PONE

Experiment 5 set out to replicate the PONE and to investigate its relationship to participants' responses, employing the same basic task as the explicit recognition conditions in Experiments 1 to 4 in which participants are asked to learn a list of words and then state whether items on a second list are old or new. If, like the ERP old/new effect, the PONE is not under voluntary control, pupil-size should increase for old items compared to new items even when participants feign amnesia and pretend not to recognise learned stimuli in the "malingering" condition (e.g. falsely respond "new" to items that are actually "old"). As pupil-size has been shown to increase in relation to cognitive load (see Chapter 1, section 1.2.2.1), and it is generally assumed that for most participants deception involves more cognitive effort than telling the truth, a third "single response" condition was included, in which participants answered "new" to all items. It was predicted that pupil-size would also increase for old items compared to new items in this condition.

5.2.1. Method

5.2.1.1. *Participants*

Twenty-six participants (6 male; age range: 19.5-30.3, $M = 23.1$, $SD = 3.3$), with normal or corrected-to-normal vision in at least one eye were recruited from the student psychology participation pool at the University of Sussex, and through personal contact. Participants were briefed with a detailed information sheet and verbal description of the task, and invited to ask questions. Written consent was obtained prior to testing and participants were fully debriefed. The experiment was approved by the relevant ethics committee.

5.2.1.2. *Materials/Apparatus*

Three study lists were created for the learning phase, each list comprising 40 nouns selected from the MRC Psycholinguistic Database. For the recognition test, three lists were constructed, each containing the 40 items that were on the corresponding study list and 40 new nouns that were not. All nouns were 7 letters long, and the old and new items were matched for familiarity and imageability, according to the K-F norms (familiarity range = 301-646, $M = 493$; imageability range = 261-630, $M = 497$). The three parallel sets of study lists and recognition tests formed blocks A to C, and were presented in black 20pt Mono-spaced font on a light grey background under fixed illumination. Words were presented using Experiment Builder software (SR-Research, Ontario) on a 21" CRT monitor. Participants viewed the monitor from a distance of 70cm and the visual angle subtended by the words was approximately 3°. Eye movements were recorded with an EyeLink II (SR-Research, Ontario), with a sampling rate of 500Hz. All items are presented in Appendix D.

5.2.1.3. *Design and Procedure*

In a within-subject design each participant completed three separate recognition memory tests. Each test contained a learning phase and a recognition phase but the instructions given differed across the tests. At the start of the experiment, participants were asked to imagine that they had recently been involved in a car accident and as a result were unconscious for 15 minutes and had to spend one night in hospital for observation. They were told that their condition had gradually improved over the following months and they had now made a full recovery. They were asked to imagine that the purpose of the test that they were about to undertake was to determine whether the accident had produced any long-term memory-impairments due to brain damage. This scenario was adapted from van Hooff et al. (2009).

In the standard instructions condition participants were asked to perform as accurately as possible during the recognition test. In the malingering condition, participants were asked to produce responses that would convince an examiner that they still had a memory-impairment. They were advised their responses should be presented in a “believable” manner, and major exaggerations, such as not remembering anything, should be avoided. To simulate real-world compensation participants were told that £10 worth of book vouchers would be awarded to the individual who best managed to simulate a believable memory deficit. In the “single response” control condition, participants were instructed to simply answer “new” to all items, regardless of whether they knew them to be old or new. This condition was intended to mimic a simple strategy that might be used by people feigning amnesia, and also allowed us to rule out any potential confounding influences on pupil-size that might result from the increased cognitive effort required to generate incorrect responses in the malingering condition.

During the learning phase, 40 study list target items were presented on screen for 2000ms with 1000ms between words, and participants were asked to remember the items. During the testing phase, 80 recognition list items (40 old targets and 40 new distracters) were presented for 1750ms, each following a 250ms mask (“&&&&&”). The mask reappeared after 2000ms and participants stated whether the word was old (previously encountered in the learning phase) or new (not previously encountered). Participants were then presented with a screen prompting them to estimate their confidence in their decision with a number between 1 and 5, where 1 represented a complete guess and 5 represented total confidence. This screen was then replaced by a drift-correction dot in the centre of the screen in preparation for the next trial.

To control against list and order effects, the condition order was rotated across participants. To determine whether any effects on pupil-size differed as a function of condition order, this variable was added as a factor to all initial statistical analyses. There were no main effects of order nor did it interact with any other factors, so, for ease of interpretation, it is not included in the analyses reported in the results section. To prevent the recognition phase instructions influencing behaviour during the learning phase (i.e. participants may not have concentrated on the study items if they knew they were going to be saying “new” to all items), instructions for the recognition phase were provided after the learning phase in each condition. Old/new judgements and confidence estimates were recorded on the computer after each recognition item.

5.2.1.4. Pupil Recording

Maximum pupil-size was recorded from the right eye during each recognition period. A Pupil Dilation Ratio (PDR; see Chapter 2, section 2.1.2.1) was calculated expressing the maximum pupil-size for each 1750ms recognition trial, as a proportion of the maximum pupil-size during that trial’s 250ms baseline.

5.2.2. Results

5.2.2.1. Behavioural Data: Old/New Responses

The proportion of old responses to old and new items was calculated for standard and malingering conditions. No old responses were made in the single response control condition. A 2 (item-type: old vs. new) by 2 (condition: standard vs. malingering) repeated measures ANOVA showed a significant main effect of item-type – in general participants responded “old” more often to old items than new items ($F(1,25) = 162.5$, $MSE = 0.027$, $p < .001$, $\eta_p^2 = .87$), a significant main effect of condition – in general participants responded “old” more often in the standard condition ($F(1,25) = 4.16$,

$MSE = 0.012$, $p < .05$, $\eta_p^2 = .14$), and a significant interaction between item-type and condition ($F(1,25) = 30.95$, $MSE = 0.03$, $p < .001$, $\eta_p^2 = .55$) – participants responded “old” to old items significantly more in the standard condition ($M = 0.79$, $SD = 0.12$) than in the malingering condition ($M = 0.56$, $SD = 0.15$; $t(25) = 5.55$, $p < .001$, $r = .552$), whereas participants responded “old” to new items significantly more in the malingering condition ($M = 0.34$, $SD = 0.15$) than in the standard condition ($M = 0.20$, $SD = 0.15$; $t(25) = 3.79$, $p < .001$, $r = .365$; see Figure 5-1).

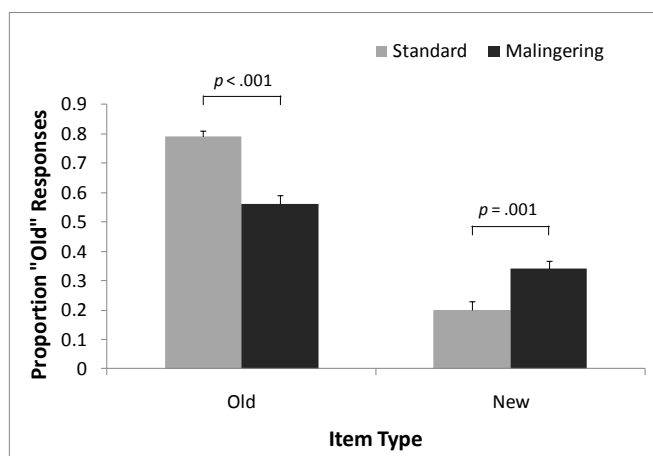


Figure 5-1: Proportion of old responses to old and new items for standard and malingering conditions. Error bars show standard error of mean.

5.2.2.2. Behavioural Data: Confidence

Confidence ratings were analysed with a 3 (condition: standard, malingering, single response) by 2 (item-type: old vs. new) repeated measures ANOVA. The main effect of item-type was significant ($F(1,25) = 27.06$, $MSE = 0.062$, $p < .001$, $\eta_p^2 = .52$) with average confidence levels for old words (3.33) higher than for new words (2.96). The main effect of condition was significant ($F(2,50) = 69.58$, $MSE = 0.523$, $p < .001$, $\eta_p^2 = .74$) with average confidence levels close to ceiling in the single response condition (4.84) and lowest in the malingering condition (3.17). Average confidence in the standard condition was 3.90. A significant condition by item-type interaction ($F(2,50)$

= 15.64, $MSE = 0.058$, $p < .001$, $\eta_p^2 = .39$) arose because confidence ratings were significantly higher for old ($M = 4.36$, $SD = 0.473$) compared to new items in the standard ($M = 3.72$, $SD = 0.745$; $t(25) = 6.30$, $p < .001$, $r = .613$) and malingering conditions (old: $M = 3.32$, $SD = 0.623$ vs. new: $M = 3.09$, $SD = 0.595$; $t(25) = 2.58$, $p < .05$, $r = .210$) but not in the single response condition (old: $M = 4.86$, $SD = 0.368$ vs. new: $M = 4.86$, $SD = 0.632$; $t(25) = 0.56$, $p > .05$, $r = .012$; see Figure 5-2).

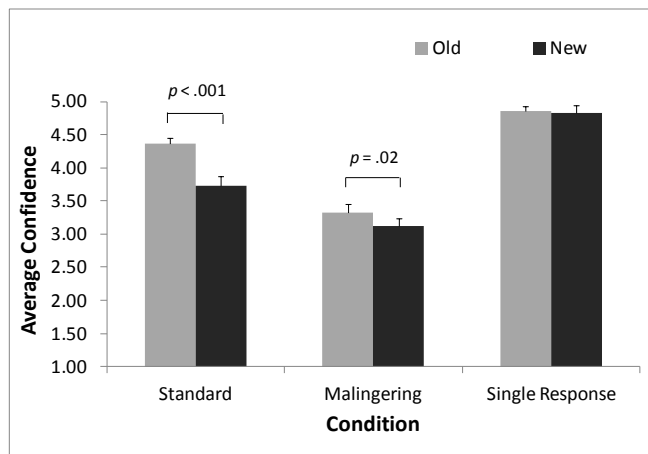


Figure 5-2: Average confidence rating for correct old and new items in each condition. Error bars show standard error of mean.

5.2.2.3. Pupil-Size Data

Average PDR for old and new items was calculated for each condition. As PDR is a function of baseline pupil-size, baseline pupil-sizes to old and new items in the three conditions were compared to ensure that any differences in PDR were not due to baseline differences. The difference was not significant ($F(1.63, 40.76) = 1.90$, $p > .05$, ns, $\eta_p^2 = .07$; Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = 6.17$, $p < .05$), therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity, $\epsilon = 0.98$).

Average PDR for old and new words in the three conditions was compared with a 2 x 3 ANOVA with item-type (old vs. new) and condition (standard vs. malingering vs.

single response) as within subject factors. There was a main effect of item-type ($F(1,25) = 47.02$, $MSE < 0.001$, $p < .001$, $\eta_p^2 = .65$) – the PDR was larger for old items compared to new items regardless of whether people were instructed to respond veridically, feign amnesia or identify all items as new. The main effect of condition was also significant ($F(2,50) = 24.37$, $MSE = 0.001$, $p < .01$, $\eta_p^2 = .49$). Average PDRs to old and new items were higher in the standard condition compared to the malingering condition and higher again in the malingering condition compared to the single response condition. These differences were significant for both old and new items (all t s > 2.6 , p s $< .05$). The main effects were, however, qualified by a significant item-type by condition interaction ($F(2,50) = 5.17$, $MSE < 0.001$, $p < .01$, $\eta_p^2 = .17$). The interaction arises because the average increase in pupil-size is smaller in the single response condition ($M = 0.009$, $SD = 0.017$) than in the standard ($M = 0.025$, $SD = 0.021$, $t(25) = 3.34$, $p < .01$, $r = .31$) or malingering conditions ($M = 0.018$, $SD = 0.021$, $t(25) = 2.07$, $p < .05$, $r = .15$) (see Figure 5-3).

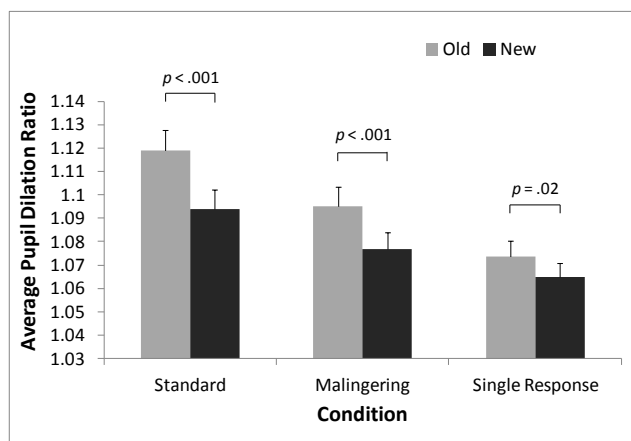


Figure 5-3: Pupil dilation ratio for old and new items in standard, malingering and single response conditions. Error bars show standard error of mean.

As participant response (old vs. new) was not meaningful in the malingering condition, it was not included as a factor in the analysis above. However, it is important to establish whether, in the standard condition, pupil-size increases for old

items that are not correctly recognised (misses), as a patient with genuine memory problems might show poor explicit recognition memory but an increase in pupil-size when targets are presented. Four participants made fewer than 5 misses, and were therefore excluded from this analysis. Average PDR to missed old items was 1.09 – the same PDR as was observed for correct rejections ($t(20) = .02$, $p > .05$, ns).

5.2.2.4. *Pupil-Size Data: Confidence Analysis*

Participants made higher confidence ratings on average to their correct “old” judgments compared to their correct “new” judgements in the standard and malingering conditions. It is important to establish the extent to which the increase in PDR that occurs when participants view old items is associated with the increase in confidence that is associated with giving an old, compared to new, response. PDR was significantly higher in the standard condition for high confidence (4 or 5; $M = 1.13$, $SD = 0.055$) compared to low confidence (< 4 ; $M = 1.10$, $SD = 0.048$) correct old judgements ($t(15) = 3.41$, $p < .01$, $r = .44$). This analysis was restricted to the 16 participants who had at least 5 high and low confidence correct old judgements, and to the standard condition because confidence judgements were not meaningful in the malingering condition (it is impossible to determine whether reduced confidence reflects a genuine uncertainty as to the correctness of their response or an understandable attempt by participants to give the impression that they have a poor memory).

Despite being overall slightly less confident in their correct rejections than their correct recognitions, participants made significant numbers of high confidence correct rejections. In order to further explore the relationship between confidence and PDR we compared PDR for correctly identified old and new items to which participants gave high confidence responses. PDR to old items that were correctly identified with

a high degree of confidence was greater than the PDR to new items that were correctly identified with high confidence for all three conditions (standard: $t(24) = 5.43$, $p < .001$, $r = .55$; malingering: $t(23) = 4.03$, $p < .001$, $r = .41$; single response: $t(25) = 2.14$, $p < .05$, $r = .15$).

5.2.3. Discussion

Experiment 5 sought to replicate the PONE and determine its relationship with participants' responses. The size of participants' pupils increased to a greater extent when they viewed old items compared to novel items in a standard recognition test; critically, this effect was also observed when participants were instructed to feign amnesia, or even just to give a "new" response to all items.

The finding that under standard recognition memory instructions, participants' relative increase in pupil-size is greater when they view old items compared to new items, replicates the findings demonstrated in experiments reported earlier in this thesis as well as previous published research (see Chapter 1, section 1.3.2.3) and demonstrates that the PONE is a robust phenomenon.

As discussed in section 1.3.2.3, it has been suggested that the PONE reflects cognitively demanding recollective processes that occur during the recognition of old items but not the correct rejection of new items (Võ et al., 2008). This interpretation builds on an extensive body of work demonstrating that increases in pupil-size occur as processing demands or cognitive load increase (see e.g., Kahneman, 1973). However, it is not clear why recognition of previously presented items should necessarily be more cognitively demanding than the correct rejection of novel items – for example, correct rejection may involve an effortful memory search, particularly when participants are using strategies such as "recall-to-reject" (Rotello, & Heit, 1999;

2000; Rotello, Macmillan, & Van Tassel, 2000) or recollection rejection (Brainerd, & Reyna, 2002; Brainerd, Wright, Reyna, & Mojardin, 2001), where they recall a similar or related item and know that the stimulus was not on the learning list (e.g., Jones, & Jacoby, 2005; Leding, & Lampinen, 2009). Studies have found that it typically takes longer to make a correct rejection than a correct recognition (e.g. Ratcliff, & Murdock, 1976).

The finding that the PONE was also observed in a single response condition (in which participants simply had to respond “new” to all items) may also seem problematic for an interpretation of the PONE based on cognitive effort – it could be argued that it takes the same amount of cognitive effort to respond “new” to a word during a recognition test when that word is old as it does when the word is new. It is possible, however, that despite the lack of any requirement for a genuine old/new decision to be made in the single response condition, recognition (and accompanying mnemonic processes) still occurred when people encountered old items. If it is assumed that it is these mnemonic processes themselves (as opposed to the cognitive effort they may involve) that are associated with the increase in pupil-size, then the present pattern of results would be expected. The PONE was greatest when participants were given standard instructions to make a genuine old/new decision for each word and diminished somewhat in the malingering and single response conditions. In the absence of any requirement to respond accurately in the malingering condition participants may have “preloaded” either an old or new response. This preloading strategy was required in the Single Response condition (and is explored further in Experiment 6). As a result in both malingering conditions less genuine recognition/recollection may have occurred, with a resulting reduction in the magnitude of the overall PONE effect when averaged across trials.

Another possibility is that the increase in pupil-size that occurs when participants view old items during a recognition test somehow reflects differences in confidence associated with correct recognition of old items compared to the correct rejection of new items. Participants did indeed give higher confidence ratings on average to their correct “old” judgments than their correct “new” judgements in both the standard and malingering conditions. If (as suggested above) the PONE reflects the operation of mnemonic processes during recognition then the extent of the pupil-size increase might be expected to be associated with confidence. Recent models of recognition memory have moved away from the idea that recollection is an ‘all or nothing’ process, instead suggesting that, like familiarity, the recollection signal may vary along a continuum (e.g., Wixted, & Stretch, 2004). If the aggregate “strength of memory” signal exceeds a certain threshold the item is identified as old (Wixted, 2007a; Wixted, & Stretch, 2004). If confidence ratings are taken as a reflection of participant’s subjective experience of the strength of this aggregate signal, and the pupil-size increase reflects the cognitive processes that drive this signal, then pupil-size increases should be greater for high compared to low confidence judgments, as was indeed the case.

Despite this relationship between confidence and successful recognition, the PONE does not simply reflect the difference in confidence between correct recognition of targets and the correct rejection of distracters. Participants can, of course, be highly confident that an item was not on the study list. When PDR was compared between only the highly confident (4 or 5) correctly identified old and new items, the PONE remained significant in all conditions. Similarly, participants were significantly more confident when making correct rejections than false alarms, but there were no differences in pupil-size. These findings suggest that whilst confidence may be

related to the magnitude of the PONE, the increase in pupil-size that occurs when participants view old items does not simply reflect a “confidence signal”.

The key finding of Experiment 5 is the demonstration that the PONE occurs even when participants are deliberately giving incorrect answers under instructions to malingering, and when they are instructed to simply identify all items as new. These results support Clark and Johnson’s (1970) finding that the PONE is not under voluntary control and show that it is independent of participants’ actual response. A similar argument has been made concerning the ERP old/new effect, and has been used to support its potential use as an index of malingering (Tardif et al., 2000; van Hooff et al., 2009). In a recent study, however, the ERP old/new effect was not observed in a group of participants instructed to malingering (Vagnini, Berry, Clark, & Jiang, 2008). Differences in procedure, in particular whether participants were asked to feign amnesia before or after learning the word list, may account for the different findings. Experiment 8 (Chapter 6) explores the potential relationship between the PONE and the ERP old/new effect.

In conclusion, this study confirms and extends previous research demonstrating that pupil-size increases more for previously encountered stimuli than for new items during a recognition memory test. Critically, this increase appears to be independent of the veracity of the behavioural responses and may have potential as a comparatively simple and easy tool with which to detect patients feigning amnesia. One thing that Experiment 5 did not address is types of malingering strategy, participants were only asked informally and retrospectively what type of strategy they used. Research has shown different strategies used in malingering including providing incomplete effort (see section 5.1.1), therefore Experiment 6 was designed to investigate the effects of the different types of malingering strategy that a

participant might use on the PONE, using three different instructions explaining how to generate poor behavioural data.

5.3. Experiment 6 – Methods of Malingering

Experiment 5 sought to determine whether the PONE was still present when participants gave deliberately incorrect responses, and found that the PONE is independent of the veracity of the behavioural response. Experiment 6 seeks to investigate different strategies that a malingering participant might use by giving specific instructions on how to generate poor behavioural data. In addition to a standard recognition condition, two conditions used strategies which were based on malingering behaviours that many neuropsychological tests were designed to detect (see section 5.1.1) – to make incomplete effort and not pay attention to stimuli during the learning phase (Incomplete Effort condition), and to answer randomly by preloading an old/new response before the item appeared on screen (Random condition).

Bradley and Janisse (1975, cited in Janisse, 1977) conducted an experiment in which participants were asked to select a numbered card, then either respond honestly (neutral condition), lie out loud to all questions by saying “no” (overt condition), or think “no” without responding (covert condition). Pupil-size was largest when responding to the critical card than the other cards in all three conditions, suggesting that due to deception and/or a response to a salient stimulus, participants were unable to hide their true response even when giving a false response or not responding out loud at all. Clark (1975, cited by Janisse, 1977) found similar results in overt and covert response conditions, and was able to detect 80% of lies based on pupil-size, compared to 85.8% using GSR and 63% using heart rate, and giving a combined accuracy of 96.7%. A fourth condition therefore required participants to not

respond behaviourally at all, but to remain silent and think “old” or “new” (Quiet condition).

In line with previous research (see Chapter 1, section 1.3.2.3) and the results of Experiments 1 to 5, it was predicted that the PONE would be present for correct responses in all conditions where a recognition decision was being made (Standard, Incomplete Effort and Quiet). This was because even when some stimuli were ignored during learning in the Incomplete Effort condition, for successfully encoded stimuli, the pupil would still reflect mnemonic processes associated with recognition. It was predicted that the PONE would not occur in the Random condition where answers were preloaded rather than a recognition decision being made. This was based on the findings of the reading conditions of Experiments 1 and 2 where no PONE occurred in the absence of a recognition decision.

5.3.1. Method

5.3.1.1. Participants

Seventy-six participants (24 male; age range: 19.4-48.0, $M = 24.5$, $SD = 5.27$), with normal or corrected-to-normal vision were recruited from the psychology course-credit and subject pools at the University of Sussex, and through personal contact.

Participants were briefed with a detailed information sheet and verbal description, and invited to ask questions. Written consent was obtained prior to testing and participants were fully debriefed. The experiment was approved by the relevant ethics committee.

5.3.1.2. *Materials/Apparatus*

One study list was created for the learning phase, comprising 40 nouns selected from the MRC Psycholinguistic Database. For the recognition test, another list was constructed, containing the 40 items that were on the study list and 40 new nouns that were not. All items were 6 letters long, matched for familiarity and imageability, according to the K-F norms (familiarity range = 436-632, $M = 556.15$; imageability range = 368-643, $M = 548.85$). The study list and recognition test were presented in black 20pt Monospaced font on a light grey background under fixed illumination. Words were presented using Experiment Builder software (SR-Research, Ontario) on a 21" CRT monitor. Participants viewed the monitor from a distance of 70cm and the visual angle subtended by the words was approximately 3° . Eye movements were recorded with an EyeLink II (SR-Research, Ontario), with a sampling rate of 500Hz. All items are presented in Appendix E.

5.3.1.3. *Design and Procedure*

In a between-subject design nineteen participants completed one of four conditions. Each condition comprised a learning phase and a recognition phase. The four conditions were standard instructions (Standard), instructions not to concentrate fully during the learning phase but perform genuinely in the recognition phase (Incomplete Effort), instructions to concentrate during the learning phase but randomly preload an old/new response before the item appeared on screen during the recognition phase (Random), and instructions to concentrate during the learning phase but to only think old/new without a behavioural response during the recognition phase (Quiet).

Prior to the start of the experiment, instructions appeared informing the participant that they would now see a list of 40 words. In the Standard, Random and Quiet conditions the instructions informed them that they would have to try and remember

these words for a later task. In the Incomplete Effort condition, the instructions stated that they should not try their best to remember the words as they were to fake a convincingly poor performance at recognition.

During the learning phase of all four conditions the same 40 study list items were presented on screen one at a time for 2000ms each in a randomised order. During the recognition phase, an instruction screen then appeared informing the participants that they were about to be presented with a list of 80 words, 40 of which they had seen before (old) and 40 of which they hadn't (new). In the Standard and Incomplete Effort conditions, participants were asked to indicate whether each word was old or new using the left and right trigger keys of the gamepad which served as the response box on the EyeLink II system. In the Random condition participants were asked to decide on their answer before the item was presented on screen, and respond with that answer using the gamepad, even if it was wrong. In the Quiet condition participants were asked to think to themselves whether the word was old or new but not to indicate their answer verbally or via the gamepad.

At the start of each trial participants saw a drift correct dot, then a mask ("&&&&&&") in the centre of the screen which lasted for 250ms. The mask was replaced by an item from the recognition list for 2000ms. The next screen asked participants to decide whether the word was old or new using a computer gamepad. This screen was then replaced by a drift-correction dot in the centre of the screen before presentation of the next trial, until all 80 items had been presented in a randomised order.

5.3.1.4. *Pupil Recording*

Maximum pupil-size was recorded from the right eye during each recognition period. A Pupil Dilation Ratio (PDR; see Chapter 2, section 2.1.2.1) was calculated expressing the maximum pupil-size for each 2000ms recognition trial as a proportion of the maximum pupil-size during that trial's 250ms baseline.

5.3.2. **Results**

5.3.2.1. *Behavioural Data*

The proportion of correct responses to old and new items was calculated for the Standard, Incomplete Effort and Random conditions (no responses were made in the Quiet condition). A 2 x 3 mixed-design ANOVA with a within-subjects factor of item-type (old vs. new) and a between-subject factor of condition (Standard vs. Incomplete Effort vs. Random) showed a significant main effect of item-type ($F(1,54) = 134.26$, $MSE = 0.018$, $p < .001$, $\eta_p^2 = .713$) – in general participants responded correctly more often to new items than to old items, a significant main effect of condition ($F(2,54) = 3.30$, $MSE = 0.014$, $p < .05$, $\eta_p^2 = .109$) – in general participants responded correctly more often in the Standard condition than the Incomplete Effort or Random conditions, and a significant interaction between item-type and condition ($F(2,54) = 27.31$, $MSE = 0.018$, $p < .001$, $\eta_p^2 = .503$). The interaction occurred because significantly more old items were correctly identified in the Standard condition ($M = 0.729$, $SD = 0.138$) than in the Incomplete Effort condition ($M = 0.571$, $SD = 0.180$; $t(36) = 3.03$, $p < .01$, $r = .204$), whereas the difference in correctly identified new items was not significant (Standard: $M = 0.787$, $SD = 0.120$; Incomplete Effort: $M = 0.734$, $SD = 0.146$; $t(36) = 1.21$, $p > .05$, ns). Additionally significantly more new items were correctly identified in the Incomplete Effort condition ($M = 0.734$, $SD = 0.146$) than in

the Random condition ($M = 0.546$, $SD = 0.069$; $t(25.7) = 5.07$, $p < .001$, $r = .500$; Levene's test indicated unequal variances ($F = 7.19$, $p = .01$), so degrees of freedom were adjusted from 36 to 25.7), whereas the difference in correctly identified old items was not significant (Incomplete Effort: $M = 0.571$, $SD = 0.180$; Random: $M = 0.512$, $SD = 0.066$; $t(22.7) = 1.35$, $p > .05$, $r = .07$; see Figure 5-4).

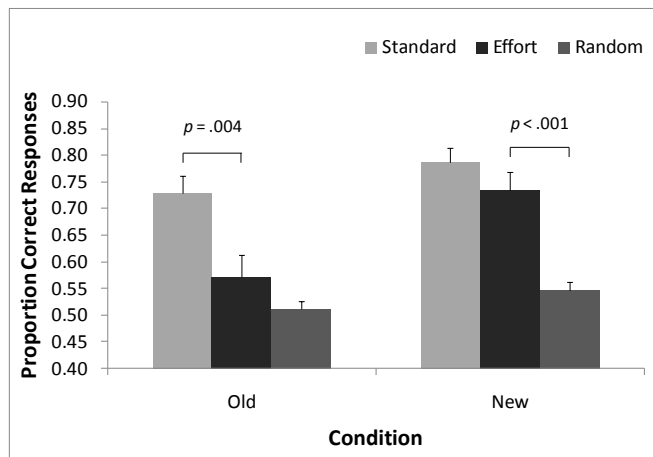


Figure 5-4: Proportion of correct responses to old and new items for Standard, Incomplete Effort and Random conditions. Error bars show standard error of mean.

5.3.2.2. Pupil-Size Data

Average PDR for old and new items was calculated for each condition. As PDR is a function of baseline pupil-size, baseline pupil-sizes for old and new items in each condition were compared to ensure that any differences in PDR were not due to baseline differences. The difference was not significant ($F(1,72) = 1.16$, $p > .05$, ns, $\eta_p^2 = .016$).

A 2 x 4 mixed-design ANOVA with a within-subjects factor of item-type (old vs. new) and a between-subject factor of condition (Standard vs. Incomplete Effort vs. Random vs. Quiet) showed a main effect of item-type ($F(1,72) = 12.76$, $MSE < 0.001$, $p < .001$, $\eta_p^2 = .151$) – the PDR was larger for old items compared to new items regardless of instructions. There was no main effect of condition ($F(3,72) = 1.83$, $MSE = 0.002$, p

$>.05$, $\eta_p^2 = .071$) and the condition by item-type interaction was not significant ($F(3,72) = 1.63$, $MSE < 0.001$, $p >.05$, $\eta_p^2 = .064$), however planned contrasts revealed that PDR was significantly larger for old items compared to new items in the Standard, Incomplete Effort and Quiet conditions (all t s > 2 , p s $\leq .05$), but was not significantly larger for old items than new items in the Random condition ($t(18) = 0.055$, $p >.05$, $r <.001$); see Figure 5-5).

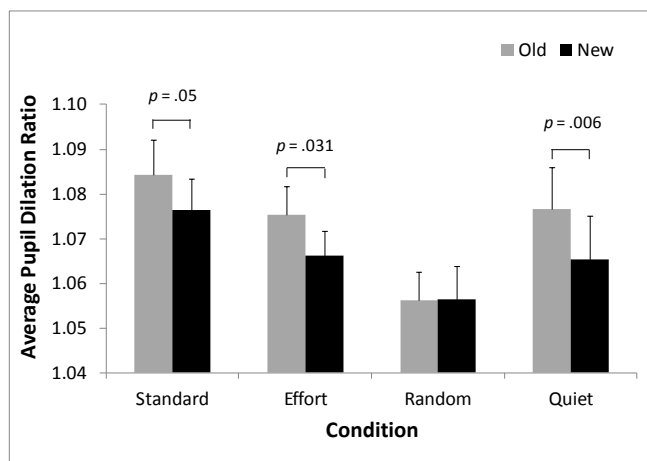


Figure 5-5: Pupil dilation ratio for old and new items in Standard, Incomplete Effort, Random and Quiet conditions. Error bars show standard error of mean.

5.3.3. Discussion

Experiment 6 sought to investigate the effect on the PONE of different malingering strategies that a participant might use, by giving specific instructions on how to generate poor behavioural data. Instructions appear to have been effective because participants in the Standard condition performed in line with previous experiments (old: 72.9%, new: 78.7%) and in the Random condition performed at chance (52.9%). Participants instructed not to learn the items in the Incomplete Effort condition performed significantly lower at correctly recognising learned old items (57.1%), whilst their performance in identifying new items was relatively intact (73.4%), as this

strategy would impair their ability to know if an item is old, but not that they have not seen an item before.

As for Experiment 5, the size of participants' pupils increased to a greater extent when they viewed old items compared to novel items under standard recognition instructions, further demonstrating the robustness of the PONE. The key finding of the present study is the demonstration that the PONE occurs even when participants are deliberately using malingering strategies. Whilst the main effect of item-type in the absence of a significant interaction could be used to statistically argue that the PONE is equivalent across conditions, suggesting that the PONE is robust to all three malingering strategies, looking at the conditions independently using planned comparisons showed that as predicted, the PONE was present in the Standard, Incomplete Effort and Quiet conditions, but was absent when participants preloaded an answer in the Random condition. The finding that the PONE exists when there is no behavioural verbal response at all is fascinating, and suggests that similar to the ERP technique, pupil-size provides a window on cognitive processes even in the absence of an overt behavioural response.

In section 5.2.3, it was suggested that in the Malingering condition of Experiment 5 participants may have "preloaded" an old or new response at random, in the absence of any requirement to respond accurately, reducing the magnitude of the overall PONE when averaged across trials. It was also suggested that this preloading strategy was a requirement in the Single Response condition (in which participants replied "new" to all items), resulting in both conditions eliciting less genuine recognition processes. However in the present experiment, explicitly asking participants to use this strategy appears to eliminate the PONE, suggesting that random preloading was not the explanation for the pattern of results in Experiment 5.

Further possible explanations for the reduction in PONE found in the Malingering and Single Response conditions of Experiment 5 are considered in the General Discussion in section 5.4.4.

As in the Random condition, the pupil-size for old and new items was also not statistically different in the long and short duration reading conditions in Experiment 2 (Chapter 3, section 3.3), when participants were asked to simply read the words out loud during the recognition phase, without the requirement to make an actual old/new recognition decision. If it is assumed that it is the mnemonic processes themselves (as opposed to the cognitive effort they may involve) that are associated with the increase in pupil-size, then we would not expect to see a difference in these conditions. It seems less likely that these processes occur unless a recognition decision is being made. Alternatively, as participants were given instructions at the start of the task, they may have employed two strategies together, firstly not paying attention during the learning phase as their knowledge was not going to be utilised, and then randomly preloading a response. While this strategy may have potential to defeat a PONE-based test of malingered amnesia, it would produce chance-level, and therefore clinically inappropriate, performance and could still identify a person who is malingering as clinical cut-offs for tests of malingered memory are well above chance (Bauer, & McCaffrey, 2006). Whilst this may catch a naïve malingerer, as discussed in section 5.1.1, the advancement of tests of malingering advance is met by improvements in strategies to defeat them.

As in Experiment 5, the two malingering conditions in which the PONE occurred, Incomplete Effort and Quiet, show an overall smaller pupil-size. Whilst it can be argued that less 'recognition' may be happening in the Incomplete Effort condition, this is not the case in the Quiet condition where participants learned items under

standard instructions and were still asked to make the recognition decision in their head. Research has shown that even the small amount of physical effort involved in making a behavioural response amplifies pupil-size (e.g., Simpson, & Paivio, 1968; see section 1.2.5), so pupil-sizes in the Quiet condition may have been smaller due to the absence of a behavioural response. In addition, because the Quiet condition had no behavioural data, it was not possible to analyse only correct trials. Including missed old items in the analysis would have reduced the average PDR for old items because misses, which can be interpreted to be items with a memory strength insufficient to reach recognition threshold, are associated with a smaller pupil-size than correctly identified old items, and statistically equivalent to that of correct rejections (see Experiment 5).

All trials were also analysed in the Random condition because ‘correct’ trials would have been an arbitrary half of the trials rather than the trials in which items were correctly recognised or rejected. Therefore, the inclusion of misses in the analysis of old items, and the inclusion of false alarms (which can be associated with an intermediate pupil-size – Otero et al., 2011) in the analysis of new items, may have masked the presence of a small PONE had it been present. Although this is not ideal, it was not possible to restrict analysis to correct items in these conditions, and it was necessary to collect data in this way for the purposes of testing potential strategies. Whilst an (extra) response could have been added to the Quiet and Random conditions, which was outside the period of the trial used for analysis of pupil-size, and asked for an accurate assessment of each item’s old/new status, this might have delayed the mnemonic processes we were attempting to measure or distracted the participant from the task. However, this idea could be developed in the future.

Building on strategies attempting to feign memory-impairment, Experiment 7 will investigate the effects of genuinely impairing memory performance in healthy participants using a divided attention task. By asking participants to perform a secondary task at encoding and retrieval, this procedure should interfere with mnemonic processing of list items, with the aim of reducing overall memory performance and observing the resulting effect on the PONE.

5.4. Experiment 7 – Emulating Memory-impairment

Experiment 5 sought to determine whether the PONE was still present when participants gave deliberately incorrect responses, and found that the PONE is independent of behavioural response. Experiment 6 sought to investigate different strategies that a malingering participant might use by giving specific instructions on how to generate poor behavioural data, and found that not paying attention during the learning phase decreased the pupil-size overall but did not diminish the magnitude of the PONE.

Experiment 7 was designed to more realistically emulate the effects of a genuine memory-impairment by dividing participants' attention at learning and recognition. Rather than asking participants to actively feign poor performance, it was hoped that reduced attention to the stimuli at encoding, retrieval, or both, would lessen genuine performance without participants having to use a strategy 'on-line' (which has its own cognitive demands and may produce slower reaction times). For example, performing a secondary task during the study phase has been shown to impair performance at recognition, however there is asymmetry in the memory system whereby divided attention at retrieval usually has only minimal effects on performance (e.g., Craik, Govoni, Naveh-Benjamin, & Anderson, 1996; Naveh-Benjamin, Craik, Guez, & Dori, 1998). It was therefore predicted that performance would be much

worse when the dual-task occurred during the learning phase than during recognition, and that the best performance would occur when no secondary task took place. As the PONE is thought to represent mnemonic processes, it was predicted that the PONE would be reduced by dividing attention at encoding during the learning task compared to performing a single task at encoding, and irrespective of whether a single or dual task was performed at recognition. This was because the memory strength for each item would be reduced due to reduced attentional resources devoted to learning during encoding and fewer deep, elaborative strategies being used. As retrieval is relatively robust to simultaneous tasks, it was predicted that dividing attention during retrieval would only reduce the PONE when attention had also been divided during the learning phase.

An effective secondary task is that of target detection (e.g., Naveh-Benjamin et al., 1998) where the participant remains alert to intermittent and irregularly timed stimuli and is asked to respond behaviourally when a target is detected. As this experiment already utilised the computer monitor and the primary task was visual, for the secondary task an auditory tone was presented. Participants were required to respond verbally when they detected a tone, whilst also either learning stimuli during the learning task or using the gamepad to indicate whether or not they recognised stimuli in the recognition task.

5.4.1. Method

5.4.1.1. Participants

Seventy-two participants (28 male; age range: 18.8-66.1, $M = 26.9$, $SD = 9.45$), with normal or corrected-to-normal vision were recruited from the psychology course-credit and subject pools at the University of Sussex, and through personal contact.

Participants were briefed with a detailed information sheet and verbal description, and invited to ask questions. Written consent was obtained prior to testing and participants were fully debriefed. The experiment was approved by the relevant ethics committee.

5.4.1.2. *Materials/Apparatus*

One study list was created for the learning phase, comprising 40 nouns selected from the MRC Psycholinguistic Database. For the recognition test, a second list was constructed, containing the 40 items that were on the study list and 40 new nouns that were not. All items were 7 letters long, matched for familiarity and imageability, according to the K-F norms (familiarity range = 293-646, $M = 491.86$; imageability range = 357-630, $M = 544.91$). The study list and recognition test were presented in black 20pt Mono-spaced font on a light grey background under fixed illumination. Words were presented using Experiment Builder software (SR-Research, Ontario) on a 21" CRT monitor. Participants viewed the monitor from a distance of 70cm and the visual angle subtended by the words was approximately 3° . Eye movements were recorded with an EyeLink II (SR-Research, Ontario), with a sampling rate of 500Hz. All items are presented in Appendix F.

A podcast of an episode of Radio 4 programme *The Archers* was downloaded from the BBC website. Using Audacity, open source software for editing audio tracks, a 100Hz tone 100ms in length was inserted into the Archers audio track on average every 5s. The positioning of the tone within a 5s bin was random, using a list of random numbers between 0 and 5 generated in Excel. The gain of the tone was -13Db in relation to the main audio track to make it difficult to detect to encourage participants to attend to it. The file was converted into mp3 format using LAME MP3

Encoder, and then played on a Sony mp3 player through speakers positioned either side of the monitor.

5.4.1.3. Design and Procedure

In a between-subject design eighteen participants completed each of four conditions: under standard instructions (Single/Single), under instructions to perform an auditory task during the learning phase (Dual/Single), under instructions to perform an auditory task during the recognition phase (Single/Dual), and under instructions to perform an auditory task during both learning and recognition phases (Dual/Dual). Each condition contained a learning phase and a recognition phase.

Prior to the start of the experiment, instructions appeared informing the participant that they would now be seeing a list of 40 words. In all conditions, instructions informed them that they would have to try and remember these words for a later task. Both the Dual/Single and Dual/Dual condition instructions proceeded to say that there would be a secondary task in which a recorded radio programme with embedded tones would be played. Participants were asked to say “tone” when the tone sounded. They were then played the first 10s of the recording so that they knew what sound to identify. During the learning phase of all conditions the same 40 study list items were presented on screen one at a time for 2000ms each in a randomised order.

During the testing phase, an instruction screen then appeared informing the participants that they were about to be presented with a list of 80 words, 40 of which they had seen before (old) and 40 of which they had not (new). In the Single/Dual and Dual/Dual conditions, participants were also informed about the secondary auditory task and those in the Single/Dual condition (who had not completed the

secondary task in the learning phase) were then given a 10s clip of the sound file. At the start of each trial participants saw a drift correct dot, then a mask (“&&&&&&”) in the centre of the screen which lasted for 250ms. The mask was replaced by an item from the recognition list for 2000ms. The next screen asked participants to decide whether the word was old or new using a computer gamepad. This screen was then replaced by a drift-correction dot in the centre of the screen before presentation of the next trial, until all 80 items had been presented in a randomised order.

5.4.1.4. Pupil Recording

Maximum pupil-size was recorded from the right eye during each recognition period. A Pupil Dilation Ratio (PDR; see Chapter 2, section 2.1.2.1) was calculated expressing the maximum pupil-size for each 2000ms recognition trial as a proportion of the maximum pupil-size during that trial's 250ms baseline.

5.4.2. Results

5.4.2.1. Behavioural Data

The proportion of correct responses to old and new items was calculated for all conditions. A 2 x 2 x 2 mixed ANOVA with a within-subjects factor of item-type (old vs. new) and between-subject factors of learning task (single vs. dual) and recognition task (single vs. dual) showed a trend effect of item-type ($F(1,68) = 3.08$, $MSE = 0.015$, $p = .08$, $\eta_p^2 = .043$) – in general participants responded correctly more often to new items than to old items (see Figure 5-6). There was also a significant main effect of learning task ($F(1,68) = 5.98$, $MSE = 0.014$, $p < .05$, $\eta_p^2 = .081$) – in general participants responded correctly more often when their learning phase contained a single task than a dual task. There was no main effect of recognition task ($F(1,68) = 0.824$, $MSE = 0.014$, $p > .05$, $\eta_p^2 = .012$).

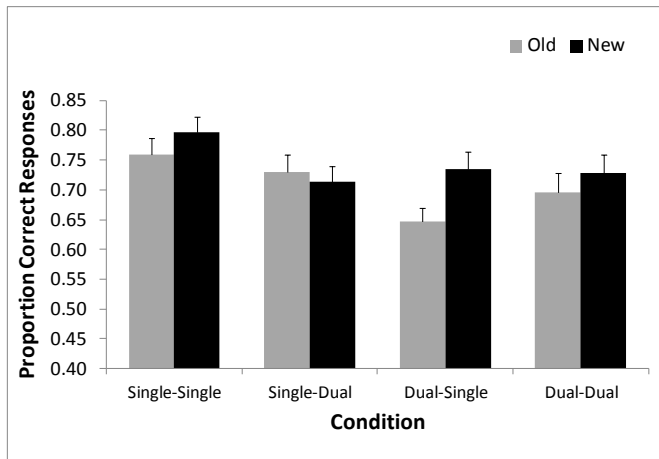


Figure 5-6: Proportion of correct responses to old and new items for all conditions. Error bars show standard error of mean.

However, the learning by recognition task interaction was significant ($F(1,68) = 3.81$, $MSE = 0.015$, $p < .05$, $\eta_p^2 = .053$) – when the recognition phase contained a single task, a dual task learning phase lead to poorer performance ($M = 0.691$, $SD = 0.068$) than a single task learning phase ($M = 0.778$, $SD = 0.079$; $t(34) = 3.56$, $p < .001$, $r = .271$). Additionally, whilst memory performance might be expected to be worst in the Dual-Dual condition, when the recognition phase contained a dual task, performance was decreased regardless of whether participants carried out a single or dual task at learning ($M = 0.722$ and $M = 0.712$, $SD = 0.105$ and $SD = 0.078$; $t(34) = 0.314$, $p > .05$, $r = .003$; see Figure 5-7).

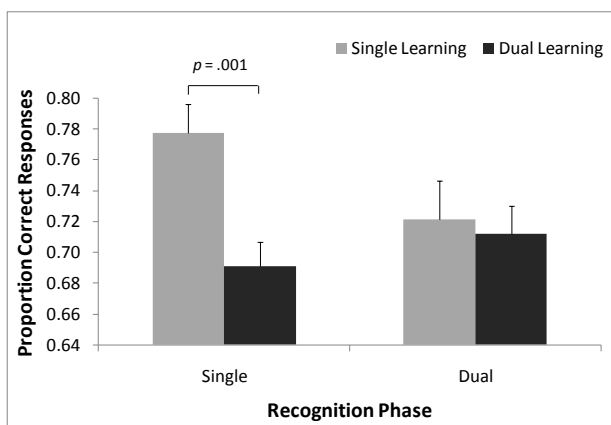


Figure 5-7: Proportion of correct responses to all items for the four combinations of single and learning task conditions. Error bars show standard error of mean.

5.4.2.2. Pupil-Size Data

Average PDR for old and new items was calculated for each condition. As PDR is a function of baseline pupil-size, baseline pupil-sizes for old and new items in each condition were compared to ensure that any differences in PDR were not due to baseline differences. The difference was not significant ($F(1,68) = 0.751, p > .05, ns, \eta_p^2 = .011$).

A 2 x 2 x 2 mixed-design ANOVA with a within-subjects factor of item-type (old vs. new) and between-subject factors of learning task (single vs. dual) and recognition task (single vs. dual) showed a significant main effect of item-type ($F(1,68) = 35.55, MSE < 0.001, p < .001, \eta_p^2 = .343$) – in general the PDR was larger for old items compared to new items. There was no main effect of learning task, no main effect of recognition task, and none of the interactions were significant (all $F_s < 1, p_s > .05$) – carrying out a secondary task did not affect the PONE. When the analysis was restricted to those items correctly identified, there was still only a significant main effect of item-type, no effects of learning or recognition task and no interactions (see Figure 5-8).

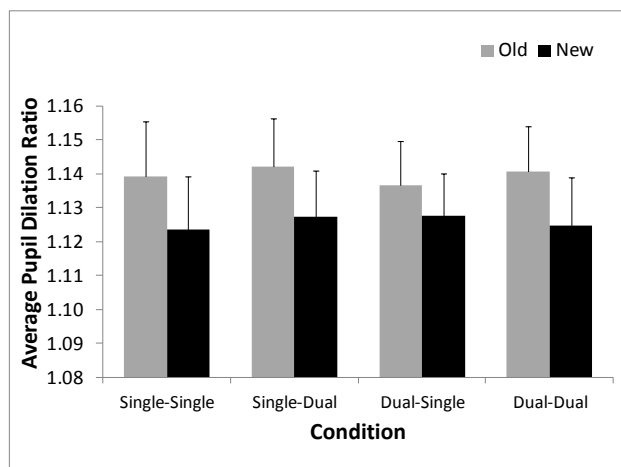


Figure 5-8: Pupil dilation ratio for old and new items in all conditions. Error bars show standard error of mean.

5.4.3. Discussion

Experiment 7 sought to determine what the effect of dividing participant attention at learning and recognition would be on the PONE. The key finding was that there was still a main effect of item-type on pupil-size, but no effect of condition, even when secondary task and learning and recognition were analysed separately. This result suggests that the PONE is a robust effect that is not diminished by dividing attention. It is less likely that the secondary task simply was not sufficiently distracting enough from the main task to impact on the PONE, since performance measures were affected by the manipulation.

It was hoped that reduced attention to stimuli at encoding, retrieval, or both, during the divided attention conditions, would lessen genuine performance relative to the Single-Single condition, without participants having to devise and apply a strategy 'on-line' (which has its own cognitive demands and may produce slower reaction times) as in Experiments 5 and 6. There was a significant main effect of learning task on performance, where participants responded correctly more often when their learning phase contained a single task than a dual task, but there was no main effect of dividing attention during recognition. This is consistent with the literature, which states that performance is reduced by divided attention at encoding but not at retrieval (Craig, Govoni, Naveh-Benjamin, & Anderson, 1996; Naveh-Benjamin, Craig, Guez, & Dori, 1998).

5.4.4. General Discussion

Experiments 5, 6 and 7 sought to further elucidate the circumstances in which the PONE occurs by manipulating participant responses and secondary task demands, and found that when a recognition decision is made, the PONE occurs even if the

behavioural response is deliberately false or absent, and is present for correct items even when a secondary task reduces genuine performance levels.

In Experiment 5 it was suggested that preloading as a strategy might explain why in the Malingering and Single Response conditions, the PONE was attenuated. It was proposed that in the absence of the requirement to give an accurate response, less genuine recognition/recollection may have occurred, reducing the overall PONE when averaged over several trials. However, in the results of the Random condition of Experiment 6, where participants were asked to preload an answer ahead of stimulus presentation as an active strategy, no PONE was found. Therefore, whilst in the Single Response condition of Experiment 5 participants were giving a predetermined answer, it is possible that they were still making an old/new judgement on some of the stimuli prior to answering “new”, and resulting in a PONE. For example, participants were still required to respond to the prompt to make a confidence judgement after each item, which may have kept participants focussed on the task as one of item recognition. In the Random condition of Experiment 6, however, participants concentrated on generating a random old/new response in the time before the next stimulus was presented, and attention may have been focussed on this task rather than the stimulus on the screen.

In the Malingering condition of Experiment 5, it would have been necessary for participants to decide themselves whether the item was old or new in order to ensure that they gave a performance that was below their best but above chance (as per the instructions for “believable” feigned memory-impairment). Therefore, some trials with new items, but which the participants identified as “old”, would have been averaged into the PDR for old items, making it smaller, and vice versa with new items becoming larger – the result being an attenuated PONE. The Random condition in Experiment

6 suggests that preloading prevents the PONE occurring, and similar to the reading conditions in Experiment 2, may be due to the absence of a requirement to make an explicit recognition decision on presented stimuli.

A notable result was that in Experiment 7, reducing the attention paid to items at encoding and retrieval affected behavioural measures but not the PONE, in a similar manner to the Incomplete Effort condition of Experiment 6, where participants were asked non-specifically to pay less attention to stimuli during the learning phase but still demonstrated a PONE. Whereas more new items are usually correctly identified than old items, in the dividing attention tasks this was not the case, similarly in the short duration recognition condition of Experiment 2, equivalent numbers of old and new items were correctly identified. This might suggest that increasing task difficulty has more of an effect on the processes involved in correct rejection than on those involved in correct recognition.

The finding in Experiment 5 that the PONE can be reliably detected even when participants are feigning amnesia and are reporting that they believe the items to be new, or when they are remaining silent as in the Quiet condition of Experiment 6, might have implications for individuals and organisations who administer neuropsychological recognition memory tests in clinical or forensic settings. These findings are similar to those of Tardif et al. (2000) who demonstrated an intact ERP old/new effect in participants asked to feign amnesia. The absence of a significant difference in PDR between old items missed, and correct rejections of new items, in the standard condition, suggests that if a patient with legitimate memory problems makes a genuine miss they would not be incorrectly identified as a malingerer on the basis of their pupil-size. Clearly it will be important to establish how pupil-size changes in genuinely memory-impaired populations when they perform this type of

recognition memory test. Laeng et al. (2007) recently investigated the pupil old/new effect in three patients with amnesia resulting from hippocampal lesions. They found that a larger pupil response occurs for new words compared to old words in these patients, similar to the findings of Experiment 1.

Given more time it would have been interesting to replicate Experiments 6 and 7 using a within-subject design to give them more statistical power and further draw out these effects. It would be interesting to look at how the PONE responds under these experimental conditions in genuinely memory-impaired populations, to further elucidate the contexts in which the PONE occurs. For example whether the PONE for items that amnesic participants correctly identify is of the same magnitude as in healthy participants, just occurring to fewer items as in the Incomplete Effort condition of Experiment 6 and the dual task conditions of Experiment 7, or whether the overall character of the PONE is diminished. Determining these parameters would indicate whether a PONE by itself implies intact recognition memory for learned items, and therefore whether the presence of a PONE in the absence of a correct behavioural response indicates malingering.

It would also be important to establish whether the PONE can be diminished by countermeasures other than the random preloading of responses seen in the Random condition of Experiment 6. Techniques have been used by Ekman et al. (2008a; 2008b) to train participants to increase and decrease the size of their pupil with the aim of using this to control aspects of a computer game, including holding their breath, hurting themselves, thinking about an emotional event, performing mental arithmetic or changing their point of focus. As so many psychological and physical events cause pupil dilation, it is possible that someone could train

themselves to, for example, perform an effortful cognitive task only when responding to new items, in order to increase pupil-size to that evoked by new items.

Although technology has become more sophisticated (and more complicated), the question remains of whether society is any better equipped to identify when someone is lying (Wolpe, Foster and Langleben, 2005). Drob (2004, cited in Ford, 2006) considers that almost any finding from current lie detection techniques could be accounted for “by something other than lying or deception” (p. 169) – current techniques are not definitive and, on their own, should not be taken as proof of lying (Ford, 2006). Most psychophysiological techniques, including pupil-size measures, require data to be averaged across multiple trials, increasing the signal to noise ratio, but also increasing the costs and time involved, making it difficult assess individual responses. An interesting extension of the experiments reported here would be to adapt the design to perform a classification analysis using bootstrapping comparison data to attempt to identify participants who are feigning memory-impairment. Performed using a sample including genuinely memory-impaired participants, participants simulating memory-impairment, and healthy controls, this would allow cut-off scores for performance and pupil-size to be established for the three categories of participants (e.g., Rogers, & Bender, 2003).

6. ERP and Pupil Old/New Effects

A Comparison

Previous chapters demonstrated that the PONE is a robust phenomenon that accompanies explicit recognition decisions, whereby participants' pupils are larger when correctly judging old items than new items. In this chapter, Experiments 8 and 9 try to establish whether there is a relationship between the PONE and another psychophysiological index of recognition memory, the ERP old/new effect.

6.1. Introduction

6.1.1. Background to ERPs and Recognition Memory

For nearly 100 years researchers have known that external events produce measurable electrical changes in the brain, with the first unambiguous experiments being conducted in the 1930s and the development and proliferation of modern Electroencephalographic (EEG) based Event-Related Potential (ERP) techniques from the 1960s onwards (Luck, 2005). Small electrical voltage differences (relative to a reference electrode) produced by the neurons of the brain are measured by scalp electrodes whilst a participant carries out a task, and amplified, digitised and stored on a computer. Whilst on individual trials consistent activity may not be visible within the continuous EEG recording, when stimulus-linked sections are averaged over a large number of trials, an ERP signal representing consistent neural activation can be distinguished from random background noise (Luck, 2005). Individual ERPs are identified as positive or negative deflections of the EEG voltage, conventionally named after their polarity, and either the approximate time at which they peak, or their

ordinal position (for example the first positive deflection is P1, the second P2 and so on; Luck, 2005).

Data from ERP recording has poor spatial sensitivity, due to the way that electricity spreads out through the conductive medium of the brain (Luck, 2005). When meeting the skull, which has a high electrical resistance, activity spreads sideways to reach the point of least resistance (Luck, 2005). The local voltage recorded by the electrode may relate to activity occurring in a distant part of the brain (Luck, 2005). In addition, the mathematical 'inverse problem' means that for a given voltage distribution it is not possible to definitively determine the sources of the underlying activity (generators; Helmholtz, 1853; Nunez, 1981; Plonsey, 1963). For these reasons ERPs alone cannot be used with confidence to localise cognitive processes; instead experiments should be designed to play to the strengths of the ERP technique (Luck, 2005). ERPs have millisecond time-resolution and can help to determine the time-course of neural activation in response to cognitive activity (Handy, 2004; Luck, 2005; van Hooff, Brunia, & Allen, 1996; see Chapter 2 section 2.2 for information about the recording, processing and analysis techniques used in this thesis). The excellent temporal resolution of ERPs complements poor temporal/good spatial resolution techniques which rely on slower metabolic processes, such as glucose uptake in PET (see Bailey, Townsend, Valk, & Maisey, 2005), or blood flow in fMRI (see Huettel, Song, & McCarthy, 2004) to localise neural activity with millimetre spatial-resolution.

During ERP data collection, participants are able to sit up to complete tasks in a more realistic situation than may be possible with other forms of neuroimaging where participants must lie horizontally in a scanner. Additionally, unlike behavioural measures, ERPs are measured directly from the scalp, and can be utilised with

participants who are unable to speak or press buttons in response to stimuli, such as young children, and in tasks where the process of interest is not measurable behaviourally, such as aspects of language processing (Luck, 2005). For well-characterised ERPs, carefully designed studies can help determine which processing stage(s) are influenced by an experimental manipulation (Luck, 2005).

Since the 1970s researchers have recorded ERPs that accompany the recognition of a previously learned item (see Donaldson, Allan, & Wilding, 2002; Fabiani, Gratton, & Coles, 2000; Friedman, & Johnson, 2000; Johnson, 1995; Rugg, & Allan, 1999; 2000, for reviews). Using the old/new paradigm, different patterns of brain activity have been observed for items recognised as old, compared to unseen new items. Specifically, correctly identified old items tend to evoke a more positive-going ERP occurring approximately 300-800ms post-stimulus onset compared to new items, misses and false alarms (Karis et al., 1984; Sanquist et al., 1980). This shift, sometimes referred to as a recognition positivity, is larger for better remembered items (Smith, 1993) and occurs later than priming positivity (a broad positivity from 250-700ms which occurs in response to repeated stimuli; e.g., Bentin & Peled, 1990; Rugg et al., 1994; Rugg, & Nagy, 1989), leading some researchers to interpret it as a confidence-related enhancement of P3 (also known as P300) – a ubiquitous positive going ERP which responds to a variety of task manipulations and overlaps spatially and temporally with memory ERP effects (Johnson, 1986; Rugg, & Nagy, 1989; Rugg et al., 1994). However other researchers have demonstrated that the old/new effect and P3 are differently affected by manipulations such as probability and previous exposure (Smith, & Guster, 1993), and when confidence is held constant the ERP old/new effect is enhanced by factors such as low word frequency because infrequent words are better remembered than common words (Rugg et al., 1995).

More recent research suggests there are two ERP old/new effects, the first concerns a frontal N400 (or FN400) wave occurring 300-500ms after stimulus presentation that is more negative for new items than old items (Wiese, & Daum, 2006), and which is also known as the MTL-N4 (Smith, Stapleton, & Halgren, 1986), medial frontal (Friedman, & Johnson, 2000), early frontal (Mecklinger, 2000) or mid-frontal old/new effect (Tsivilis, Otten, & Rugg, 2001; Curran et al., 2006). This frontal old/new effect bears similarities to the N400 evoked by visual or auditory word stimuli in the semantic processing literature (Kutas, & Hillyard, 1980), but differs functionally and topographically (Curran, Tucker, Kutas, & Posner, 1993; Curran et al., 2001), and is a sensitive index of the degree of mismatch between a word and a previously established semantic context (semantic priming; e.g., Bentin, & McCarthy, 1994; Bentin, McCarthy, & Wood, 1985; Holcomb, 1998). The N400 responds to word frequency (which affects familiarity), being larger for less frequent words (Van Petten, & Kutas, 1990; 1991), and to stimulus repetition (Van Petten, Kutas, Kluender, Mitchiner, & McIsaac, 1991), which may explain its similarity to the frontal old/new effect, where new words are a mismatch with the experimental learning context. However, the FN400 effect has also been reported for pictures (Curran, & Cleary, 2003), faces (Nessler, Mecklinger, & Penney, 2005) and objects (Mecklinger, von Cramon, & Matthes-von Cramon, 1998).

The second old/new effect appears as a parietal positive-going wave from around 400-800ms, that is more positive for old items than new items (see Johnson, 1995, for a review; Allan, Wilding, & Rugg, 1998; Friedman, & Johnson, 2000; Mecklinger, 2000; Rugg, 1995; Wilding, & Sharpe, 2003). As is common in ERP research (Luck, 2005), this component overlaps with the P300 component (Bentin, & McCarthy, 1994; Spencer, Vila Abad, & Donchin, 2000; ERP waveform shapes reflect the sum of underlying positive and negative going latent ERP components, which may be

independent yet difficult to isolate, leading to issues with interpretation – see Luck, 2005), and is referred to as the P300 old/new difference (Johnson, 1995), the parietal old/new effect (Allan et al., 1998), the late ERP old/new effect (Rugg, 1995), the MTL-P3 (Smith et al., 1986), the Late Positive Complex (LPC) (Olichney et al., 2000), and the P600 old/new when studying sentences (Rugg, & Doyle, 1992; Curran, 1999; Curran et al., 2006).

The parietal LPC old/new effect is suggested to index recollective processes (see Allan et al., 1998, for a review; Friedman, & Johnson, 2000; Mecklinger, 2000; Wilding, & Sharpe, 2003), whereas the FN400 old/new effect is thought to index familiarity (Rugg et al., 1998a). These two old/new effects have been used to provide electrophysiological evidence in support of dual-process models of recognition memory, owing to the fact that they respond differently to experimental manipulations designed to differentiate recollection and familiarity. For example, when using a remember/know paradigm, “remember” responses produce a larger parietal old/new effect than “know” responses (Curran, 2004; Düzel et al., 1997; Friedman, 2004; Rugg et al., 1998b; Smith, 1993; Trott et al., 1999). Functional imaging has also shown different patterns of brain activation associated with recollection and familiarity, which are congruent with ERP topography (e.g., Wheeler, & Buckner, 2004).

Consistent with its association with recollective processes, the LPC old/new effect has been associated with the retrieval of contextual “source” information about the original learning experience, such as temporal source – whether an item was on the first or second of two learning lists (Trott, Friedman, Ritter, & Fabiani, 1997), voice – whether the speaker of an auditory word stimulus was male or female (Rugg et al., 1998b; Wilding, & Rugg, 1996; 1997a), or stimulus modality – whether the item was read or heard (Wilding, Doyle, & Rugg, 1995; Wilding, & Rugg, 1997b).

Levels of processing manipulations, which are believed to affect recollection more than familiarity, have a greater effect on the parietal old/new effect – which shows enhanced positivity for correctly identified deeply encoded words than for shallowly encoded words (Paller, & Kutas, 1992; Paller, Kutas, & McIsaac, 1995; Rugg et al., 1995) – than on the mid-frontal old/new effect, which does not differentiate between deeply and shallowly encoded items (Rugg et al., 1998a). Similarly, the LPC old/new effect is larger for old and new words than for old and new pseudo-words, however the mid-frontal old/new effect is similar for old and new words and old and new pseudo-words (Curran, 1999). Instead, the mid-frontal old/new effect is more negative for old pseudo-words than for old words, possibly due to its sensitivity to contextual mismatch (Curran 1999).

Curran (2000) used a plurality recognition task (see Chapter 1, section 1.3.1.2 for a description of this manipulation) to differentiate familiarity and recollection in an ERP study. Participants completed three blocks, in each of which they learned 40 words and were tested on 60 words (20 old, 20 new and 20 similar lures of reversed plurality). As expected, participant responses showed a higher rate of false alarms to lures than to new items due to increased familiarity. The mid-frontal old/new effect at 300-500ms was more negative-going for correctly identified new items than correctly identified old items and false alarms to similar lures, reflecting a difference in familiarity, whereas the parietal LPC at 400-800ms was larger for old items than for new or similar lures, reflecting recollection. Based on the qualitative difference in topographical distribution of difference waves for familiarity (similar-new) and recollection (old-new) in the two time windows, Curran (2000) concluded that their findings produced strong evidence of dissociation between familiarity and recollection. This pattern of results was replicated for old, new and reversed picture stimuli among participants who performed well at discriminating between old and similar word items

(Curran, & Clearly, 2003). Participants who were poor at discriminating word stimuli showed the same frontal familiarity effect, but did not differentiate between correct old items and false alarms at parietal electrodes (Curran, & Clearly, 2003).

However, some researchers have shown posterior old/new effects associated with familiarity and anterior old/new effects associated with recollection, when using previously unknown faces (MacKenzie, & Donaldson, 2007; Yovel, & Paller, 2004). Yovel and Paller (2004) used faces rather than word stimuli because they observe that words have pre-existing levels of familiarity, which are a potential confound and mean that learned items are not compared with truly novel stimuli. By pairing unknown faces with spoken occupations during the learning phase, they were able to isolate recollection and familiarity by asking participants to qualify an “old” judgement with whether or not occupation, or other detail, could also be retrieved (indicating recollection), or whether no additional detail could be retrieved (indicating familiarity; Yovel, & Paller, 2004). As none of the faces had been previously seen, there could be no pre-existing familiarity or recollection.

Recollected items evoked a parietal old/new effect in the 500-700ms time window, consistent with previous studies (e.g., Curran, 2004; Düzel et al., 1997; Friedman, 2004; Rugg et al., 1998b; Smith, 1993; Trott et al., 1999). Familiar items also demonstrated a parietal old/new effect between 500-700ms, but no early mid-frontal old/new effect (Yovel, & Paller, 2004). Yovel and Paller (2004) somewhat controversially suggest that for faces, familiarity is represented by the parietal old/new effect, rather than the frontal old/new effect, which they attribute to the ‘conceptual priming’ associated with the use of lexical stimuli (MacKenzie, & Donaldson, 2007). In addition, the topographical distribution of the old/new effects for recollection and familiarity were statistically equivalent and they concluded a shared set of underlying

neural generators for both types of recognition judgement (Yovel, & Paller, 2004), which could be interpreted to support single-process models of recognition memory (MacKenzie, & Donaldson, 2007). However, MacKenzie and Donaldson (2007) extended this finding using faces devoid of hair, ears or background, paired with names, in shorter test blocks and recording using a higher density of electrodes (61 vs. 21 scalp locations). In addition to Yovel and Paller's (2004) posterior familiarity old/new effect, they found a novel anterior recollection old/new effect, demonstrating dissociation of recollection and familiarity by means of distinct ERP components (MacKenzie, & Donaldson, 2007). The authors interpreted their findings from a dual-process perspective and suggested that recognition old/new effects may differ under different experimental settings.

6.1.2. Pupil Responses and ERPs

As seen in Chapter 1, section 1.2.6, some researchers have recorded concurrent pupil-size with other psychophysiological measures. Interestingly, some pupil responses appear to have specific parallels in the ERP literature, for example in vigilance or reaction time tasks, a pupil dilation has been observed beginning 1000-1500ms prior to the presentation of an expected stimulus or the requirement to make a response, and which varies with the force of anticipated movement (e.g., Bradshaw, 1969; Klingner, 2010; Richer, Silverman, & Beatty, 1983). This response dilation has been likened to the ERP components known as Contingent Negative Variation (CNV; Richer, & Beatty, 1985; Richer et al., 1983; Rohrbaugh, Syndulko, & Lindsley, 1976) and the Lateralised Readiness Potential (LRP; Becker, Iwase, Jürgens, & Kornhuber, 1976; Luck, 2005), both of which precede responses by around 1000-1500ms and may represent pre-motor preparation to react to a stimulus (Beatty, & Lucero-Wagoner, 2000). Response preparation has also been shown to occur in other

psychophysiological measures such as heart rate (Coles, & Strayer, 1985) and electromyogram (Brunia, & Vingerhoets, 1980).

Other researchers have shown that stimulus probability is inversely related to the size of both task-evoked pupil dilations and the P300, whereby rare, low-frequency or unlikely stimuli evoke the largest pupil dilations and P300 amplitudes (Bock, 1976, cited by Janisse, 1977; Friedman et al., 1973; Qiyuan, Richer, Wagoner, & Beatty, 1985; Steinhauer, 1982). Some of these studies have looked at ERP and pupil-size measures simultaneously. For example, after observing that P300 and pupil dilation behaved in a similar manner in response to stimulus probability, Friedman et al. (1973) measured them concurrently, and found that both the P300 and pupil-size were inversely and monotonically related to stimulus probability in a guessing game. Steinhauer (1982) found that the P300 and pupil-size both increased in relation to bet value, event uncertainty and the absence of expected feedback in a gambling task, and were much larger when participants selected bets rather than when the computer chose for them.

Just et al. (2003) concluded that the correspondence between ERPs, pupillometry, and also fMRI responses, to the same cognitive tasks, indicate a common underlying “construct”, which they believe to be cognitive load (see Chapter 1, section 1.2.2.1). Nieuwenhuis, Aston-Jones and Cohen (2005a; Nieuwenhuis et al., 2011a) proposed that the P300 corresponds to the neuromodulatory Locus Coeruleus Norepinephrine (LC-NE) system, reacting to perceptual decision-making in stimulus evaluation (the role of the LC-NE system in stimulus evoked pupil-size change was discussed in Chapter 1, section 1.1.2.7). Several researchers have recently explored this model, utilising links between pupil-size, LC activity and task exploitation (e.g., Gilzenrat et al., 2010; Jepma, & Nieuwenhuis, 2011; Murphy et al., 2011).

Murphy et al. (2011) used an extended auditory oddball paradigm to see whether, on a trial by trial basis, P300 also indexed fluctuations in task performance predicted by Adaptive Gain Theory (AGT; Aston-Jones, & Cohen, 2005; see Chapter 1, section 1.2.4.1), and how it related to tonic and phasic changes in pupil diameter. Twenty-four participants were asked to respond to 1000Hz target tones presented on 20% of trials and ignore the remaining 500Hz standard tones presented the other 80% of the time. They found that both P300 and pupil-size reflected changes in task engagement as described by AGT. On trials with an intermediate prestimulus pupil-size, and large stimulus-evoked pupil dilations, P300 amplitudes were found to be large, and task performance better, than when prestimulus pupil-size was larger or smaller, and stimulus-evoked dilations were smaller. This pupillary behaviour was assumed to reflect intermediate tonic LC activity interspersed with phasic bursts of LC activity, consistent with the operation of the “phasic” LC mode thought to promote task engagement. Murphy et al. (2011) concluded that, in addition to pupil-size, the P300 may also index LC exploration/exploitation mode.

Kuipers and Thierry (2011) recorded concurrent pupil-size and ERPs to investigate the relationship between semantic integration, reflected by the N400 component, and phasic pupil dilation influenced by the LC. Maximal pupil-size usually occurred at least 1000ms after stimulus presentation onset (Beatty, 1982b), however, Steinhauer and Hakerem (1992) observed an initial peak dilation beginning 200ms after stimulus onset, reaching maximum amplitude between 500–600ms, only slightly later than the N400. Kuipers and Thierry (2011) investigated this early pupil peak in conjunction with ERPs by presenting participants with semantic matching/non-matching spoken word-picture and picture-spoken word pairs, and asking them to passively attend rather than to engage in a task. In the word-picture condition, the N400 amplitude was larger for matching than non-matching pairs, and pupil dilations were larger for

non-matching pairs than matching pairs. In the picture-word condition, the N400 showed larger amplitude for matching than non-matching pairs, but pupil dilations did not differ between conditions. Despite these findings, the authors focussed on spurious positive correlations between pupil-size and ERP waveforms at 16ms intervals throughout an 850ms epoch, and interpret results from the entire window in relation to the N400. Later significant differences in pupil-size (from 366ms onward) in the word-picture condition extended to the end of the 850ms epoch (and likely extended beyond it) but were not analysed.

The interpretation of Kuipers and Thierry's (2011) data is difficult. As is clear from the reported waveforms, there were large differences between the conditions before even comparing matching vs. non-matching item pairs, but Kuipers and Thierry (2011) performed separate ANOVAs for each condition disallowing any test for a significant effect of condition. The analysis also failed to provide a sense of the topographical distribution of the effects found by the authors. Major light-reflex confounds were introduced because display brightness decreased from high to low in the word-picture condition and increased from low to high in the picture-word condition. Additionally the pupil data in the picture-word condition were incorrectly baselined. There was also a more subtle confound of stimulus repetition (each pair was repeated twice in each condition) which was not included as a factor in the analyses – repeated mismatched pairs might be more memorable than matching pairs, and therefore be encoded more strongly (e.g., Otero et al., 2011), thus leading to a larger pupil dilation to non-matching pairs than matching pairs; alternatively these items may cause repetition suppression (e.g., Schacter, & Buckner, 1998), leading to a smaller pupil dilation. The authors also over-interpreted their findings in line with LC-NE influences on task performance, stating that there is no functional connection between the auditory orienting response and pupil dilation, despite acknowledging that auditory

neurons are known to respond to NE (only released by the LC) in the monkey (Foote et al., 1975) and that pupil dilation can be triggered by auditory input (Beatty, 1982a). They also ignored any possible link between small pupil dilation and decreased phasic LC firing due to task disengagement, when there was no task, and participants were asked to passively look at the screen (Gilzenrat et al., 2010).

The role of the N400 as a sensitive measure of context mismatch was also omitted. Instead the authors concluded that changes in pupil-size in their study were due to accommodation, and that decreased phasic LC firing increased the “effort” involved in semantic integration (as measured by the N400), which decreased pupil dilations. The continuation of this line of argument is that larger phasic LC input, which would increase semantic integration efficiency, would also reduce effort (and the N400) but increase pupil-size. We have only to look at the wealth of literature spanning the last six decades, demonstrating increases in pupil-size associated with cognitive effort, to question both the results and the conclusion (for a review see Beatty, & Lucero-Wagoner, 2000; Beatty, 1982; Granholm, & Steinhauer, 2004; Hess, & Polt, 1964; Janisse, 1977; Kahneman, 1973).

Few studies measure concurrent ERP and pupil-size. Van Droof et al. (2010) indirectly tested recognition memory for words by looking at receptive vocabulary knowledge in nonverbal autistic participants and found that peak dilation was larger for known words than unknown words and that the N400 was enhanced for mismatched known words. Stone and Rothenheber (1992) added EEG and pupillometry to traditional polygraph measures during an oddball experiment where participants were asked to count instances of a known photograph among a series of unknown photos. They concluded that: “Although these results were encouraging our

findings at this point are inconclusive thus warranting additional study” (Stone, & Rothenheber, 1992, p.73).

6.2. Experiment 8 – Strength of Memory Effect

Experiment 8 employs a novel approach to understanding recognition memory, concurrently measuring pupil-size and ERPs to study the effects of a memory strength manipulation. The procedure and design were based on an ERP study by Finnigan et al. (2002) who recorded continuous EEG whilst participants performed an old/new recognition test on items that were unstudied (new), studied once (weak) or studied three times (strong) during learning. Consistent with Van Petten et al. (1991), they found a graded FN400 component which had a more negative-going amplitude for new items than weak, and for weak than strong items. Like Yovel and Paller (2004), their early old/new effect was maximal over parietal electrode sites. They found larger amplitude of the LPC component (between 500-700ms) for strong items than weak and for weak than new items. The LPC amplitude was also larger for correct than incorrect decisions, with maximal amplitude at centro-parietal electrodes. This design was selected because of the graded effect of the memory manipulation on the psychophysiological responses, and it was hoped that it would also produce a graded pupil response in that PDR for strong items would be larger than for weak items, which would be larger than for new items. Comparable memory strength effects in the two measures would suggest they index the same underlying events and provide support for the idea that the PONE reflects mnemonic processes. The manipulation was expected to work because stimulus repetition has been shown to enhance memory performance on behavioural measures (e.g., Leding, & Lampinen, 2009; Yonelinas, 2002).

6.2.1. Method

6.2.1.1. *Participants*

Twenty-two right-handed native English speaking participants (9 male; age range: 19.3-50.6, $M = 26.0$, $SD = 1.67$), with normal or corrected-to-normal vision in at least one eye and no self-reported psychiatric or neurological conditions, were recruited from the student psychology participation pool at the University of Sussex and through personal contact. Participants were briefed with a detailed consent form and verbal description, and invited to ask questions. Written consent was obtained prior to testing and participants were fully debriefed at the end. The experiment was approved by the relevant ethics committee. Four participants failed to contribute more than thirty artefact-free correct ERP trials to all three item-types and were excluded from the analysis.

6.2.1.2. *Materials/Apparatus*

Three study lists were created for the learning phase, each list comprising 60 nouns selected from the MRC Psycholinguistic Database, half of which were included three times in the respective learning list. For the recognition test, three lists were constructed, each containing the 30 items that were presented once on the corresponding study list (“weak”), the 30 items that were presented three times on the study list (“strong”), and 30 new nouns that had not previously been seen (“new”). All items were 5 letters long, matched for frequency, familiarity and imageability, according to the K-F norms (frequency range = 10-40, $M = 20.3$; familiarity range = 351-618, $M = 515$; imageability range = 293-632, $M = 507$). The three parallel sets of study lists and recognition tests formed blocks A to C, and were presented on a computer monitor in white 20pt Monospaced font on a black background under fixed

illumination. Words were presented using E-Prime 2.0 software (Psychology Software Tools Inc, Pennsylvania) on a 17" CRT monitor, which participants viewed from a distance of 50cm and the visual angle subtended by the words was $\sim 3^\circ$. Eye movements were recorded with desk mounted EyeLink 1000 (SR-Research, Ontario), with a sampling rate of 500Hz. All items are presented in Appendix G.

The experiment took place inside a Faraday cage. Continuous EEG recordings were acquired from the scalp by a Net Amp and 128 electrode dense-array Geodesic Sensor Net (Tucker, 1993), in conjunction with Net Station software package (Electrical Geodesics Inc, Oregon), filtered online by bandpass 0.01-100Hz and digitized at a sampling rate of 500Hz. Both the EEG and eye movement recordings were triggered simultaneously by E-Prime; Net Station commands were sent via an Ethernet cable by the E-Prime Net Station extension, and EyeLink commands via a modified parallel cable and a custom E-Prime script that turned the cable pins on and off to stop and start eye-tracker recording. Messages indicating the beginning and end of each trial, and the onset and offset of stimuli presentation were also sent to both Net Station and EyeLink in order that pupil and EEG trials could be aligned. Net Station also received additional trial messages including item-type and participant responses, which were made using a button box.

6.2.1.3. Design and Procedure

In a within-subject design participants completed 3 recognition blocks under standard instructions. Each block contained a learning phase and a recognition phase. During the learning phase, 120 study list target items (30 items presented once, 30 items presented three times) were displayed on screen for 1000ms with 200ms of blank screen between words, and participants were asked to remember the items. During the recognition phase, 90 list items (30 new, 30 weak, 30 strong) followed a 500ms

fixation cross and a 1000ms mask of “HHHHH”, and were presented on screen for 1000ms before being remasked for 1000ms. Participants were then presented with a sign indicating that they could blink, and after 400ms a response prompt appeared that remained on screen until they responded (see Figure 6-1).

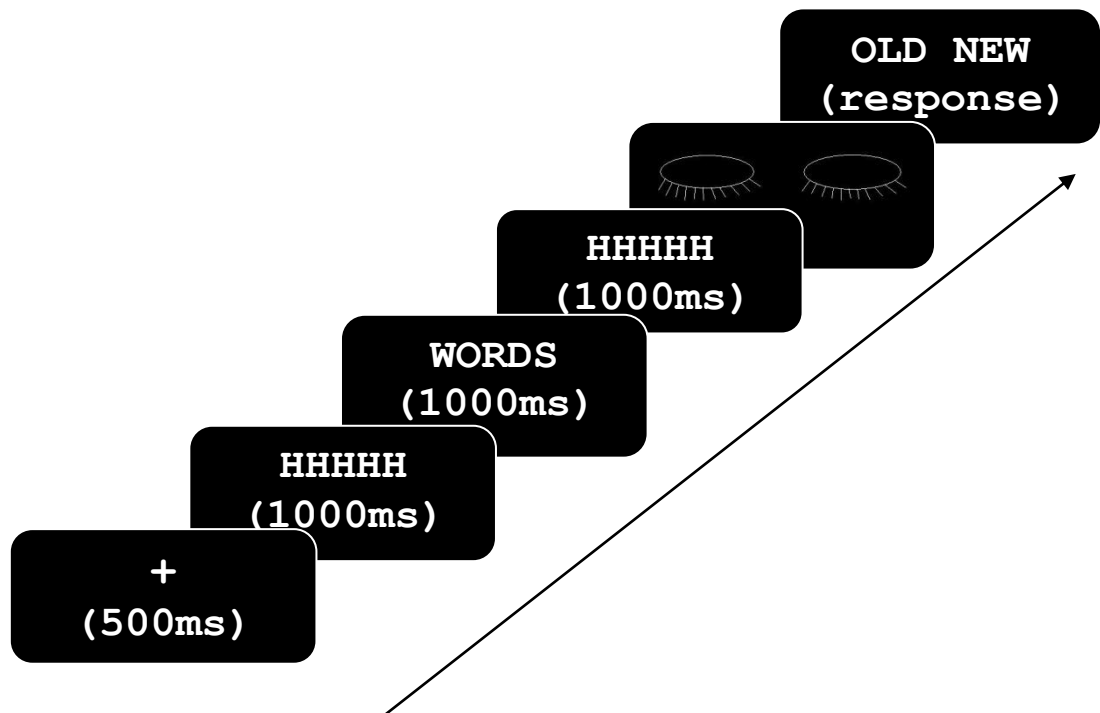


Figure 6-1: Experimental procedure.

Participants were asked to wash and brush their hair before the application of the electrodesic net to their head. Once seated in the Faraday cage they were required to use a chin rest to enable accurate eye-tracking. Participants were reminded at the start of each recognition block that their eye movements and brain waves were being recorded and to remain still and blink only when prompted by the blink screen; they were also shown the impact on the EEG traces of blinks and eye-movements. Participants were prompted to press a button to indicate whether the word was old (target previously encountered in the learning phase) or new (not previously encountered). This response screen was replaced by a fixation cross in the centre of the screen before presentation of the mask followed by the next item.

Using standard recognition instructions, participants were asked to perform as accurately as possible. In a within-subject design all participants viewed all study and recognition lists once. To control against list and order effects, items were randomised within blocks, and blocks were rotated across participants. Instructions were repeated at the beginning of each block and participants were able to take a break or initiate the next block when they felt ready with a verbal response. A response device with two buttons corresponding to “old/new” answers was provided and button configuration was counterbalanced across participants. Old/new judgements were recorded, by the computer running E-Prime, after each recognition item. Maximum pupil-size was recorded by the EyeLink host computer during the time the item was on screen during the recognition test, and EEG activity was recorded continuously by the Net Station host. Preparation and experimental procedure lasted approximately 1 hour, with the task lasting around 29 minutes.

6.2.1.4. Pupil Recording

Maximum pupil-size was recorded from the right eye during each recognition period. A Pupil Dilation Ratio (PDR; see Chapter 2, section 2.1.2.1) was calculated expressing the maximum pupil-size for each 2000ms recognition trial as a proportion of the maximum pupil-size during that trial’s 200ms baseline.

6.2.1.5. Electrophysiological Recording and Analysis

EEG was continuously recorded with a vertex reference. Vertical and horizontal eye-movements were monitored by using two bipolar ocular electrodes. Impedance was kept below 50k Ω . Sampling rate was 500Hz and an online 0.01-100Hz band-pass filter was used. Offline the continuous EEG was segmented into epochs from 200ms before to 1000ms after stimulus onset. Segments with artifacts exceeding $\pm 75\mu V$ were automatically rejected and electro-oculogram (EOG) artifacts were detected

using Net Station Waveform Tools software package (Electrical Geodesics Inc, Oregon). Manual eye artifact rejection was also used because Net Station algorithms eliminated too many trials; in total the average number of trials lost to artifact rejection per participant per condition were: 9.47 new, 4.76 weak, and 7.79 strong. Final average trial numbers per participant per condition were: 63.61 new, 46.72 weak, and 57.83 strong. Bad channels were replaced by spline-constructed data from adjacent channels using Waveform Tools for creation of topographical maps (average 1.59 channels per participant); however reconstruction was not performed for channels included in the statistical analysis. Segments were baseline-corrected over the 200ms pre-stimulus interval and re-referenced to the average mastoid electrode. Separate grand-average ERPs were computed for strong hits and misses, weak hits and misses, correct rejections and false alarms. Only correct trials were included in statistical analyses. Data were discarded from participants for whom ERP averages did not comprise at least 30 artifact-free trials (4 participants in total). Grand-averages were low-pass filtered at 40Hz prior to plotting and after statistical extraction.

6.2.2. Results

6.2.2.1. Behavioural Data

The proportions of correct responses to new, weak (presented once at learning) and strong (presented three times at learning) items were calculated and averaged 81.2% ($SD = 11.9\%$) correct new items, 57.2% ($SD = 14.3\%$) correct weak items and 72.9% ($SD = 12.0\%$) correct strong items (see Figure 6-2). A one-way repeated-measures ANOVA revealed a significant main effect of item-type ($F(1.17, 19.8) = 18.1$ $MSE = 0.025$, $p < .001$, $\eta_p^2 = .516$). Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = 20.1$, $p < .001$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = 0.583$). Bonferroni-

corrected subsidiary t-tests revealed that, in general, participants correctly identified significantly fewer weak items than strong ($t(17) = 9.65, p < .001, r = .846$) or new items ($t(17) = 4.81, p < .001, r = .577$).

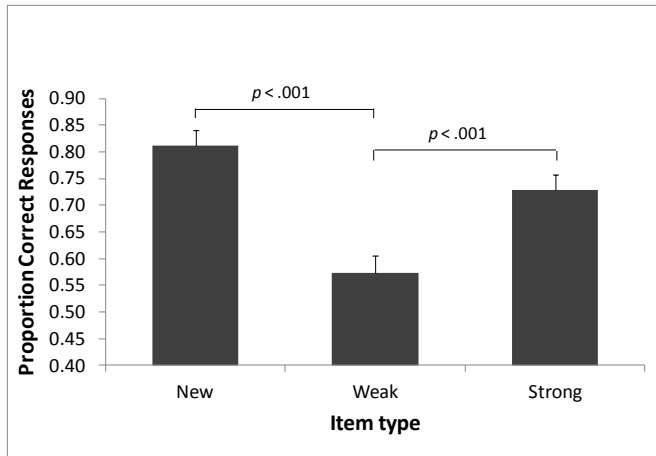


Figure 6-2: Proportion of correct responses to new, weak and strong items. Error bars show standard error of mean.

6.2.2.2. Pupil-Size Data

Average PDR was calculated for correctly identified new, weak (presented once at learning) and strong items (presented three times at learning). As PDR is a function of baseline pupil-size, baseline pupil-sizes for old and new items in each condition were compared to ensure that any differences in PDR were not due to baseline differences. The difference was not significant ($F(2,34) = 1.78, p > .05, ns, \eta_p^2 = .095$).

A one-way repeated measures ANOVA, with within-subject factor of item-type (new vs. weak vs. strong) showed a main effect of item-type ($F(2,34) = 9.02, MSE < 0.001, p < .001, \eta_p^2 = .347$). Planned t-tests revealed that in general participants' pupils dilated more to correctly identified weak old items ($M = 1.138, SD = 0.077$) than to correctly identified new items ($M = 1.118, SD = 0.0717, t(17) = 5.51, p < .001, r = .641$). Counter to predictions, participants' pupils also dilated more to correctly identified weak old items than to correctly identified strong old items ($M = 1.127, SD = 0.0817,$

$t(17) = -2.09, p < .05, r = .204$), and only approached significance for new and strong items ($t(17) = 1.78, p = .09, r = .157$; see Figure 6-3).

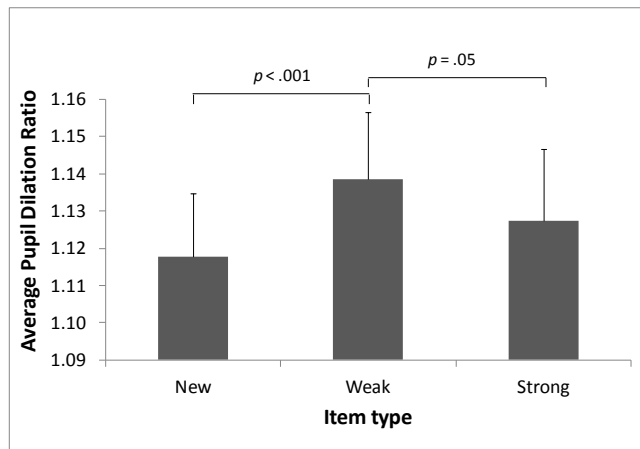


Figure 6-3: Pupil dilation ratio for correctly identified new, weak and strong items. Error bars show standard error of mean.

6.2.2.3. *Event-Related Potentials*

Figure 6-4 shows grand-average ERP waveforms at midline frontal, central and parietal electrodes for correctly identified new, weak and strong items. Waveforms diverge at around 300ms post-stimulus, and last for the rest of period of interest (the following 700ms). Inspection of the grand-average ERPs reveals a negative-going waveform between 300-500ms with a centro-parietal distribution, which is larger in amplitude for new and strong items compared to weak items, and maximal at P7. This component is similar in polarity and timing to the mid-frontal old/new effect (e.g., Mecklinger, 2000; Wiese, & Daum, 2006) but has a more posterior distribution. This is followed by a positive-going waveform between 500-700ms with a parietal distribution, which is larger in amplitude for weak and strong items compared to new items, and is also maximal at P7; this component is similar in polarity, timing and distribution to the LPC old/new effect (e.g., Johnson, 1995; Allan et al., 1998). At frontal electrodes there is also a negative-going waveform between 500-700ms which is larger in amplitude for strong items than new or weak items.

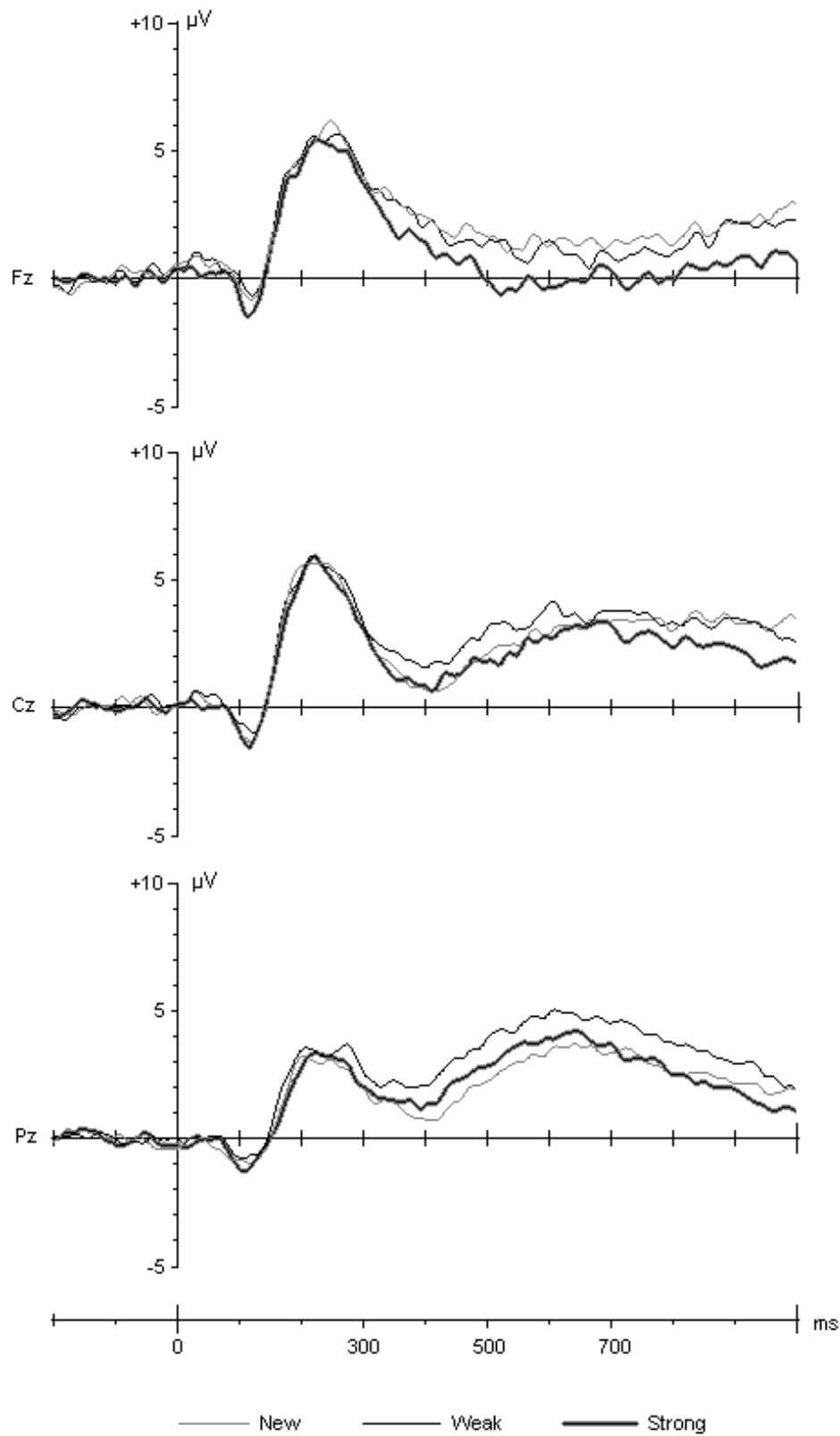


Figure 6-4: Grand average ($N = 18$) ERPs for correctly identified new, weak (presented once) and strong (presented three times) items at midline frontal, central and parietal electrodes. Mean numbers of individual ERP trials per strength condition per participant were: new: 63.61; weak: 46.72; strong: 57.83. The scale bar indicates amplitude (in μV) and time course of activity (in ms). Positive plotted upwards.

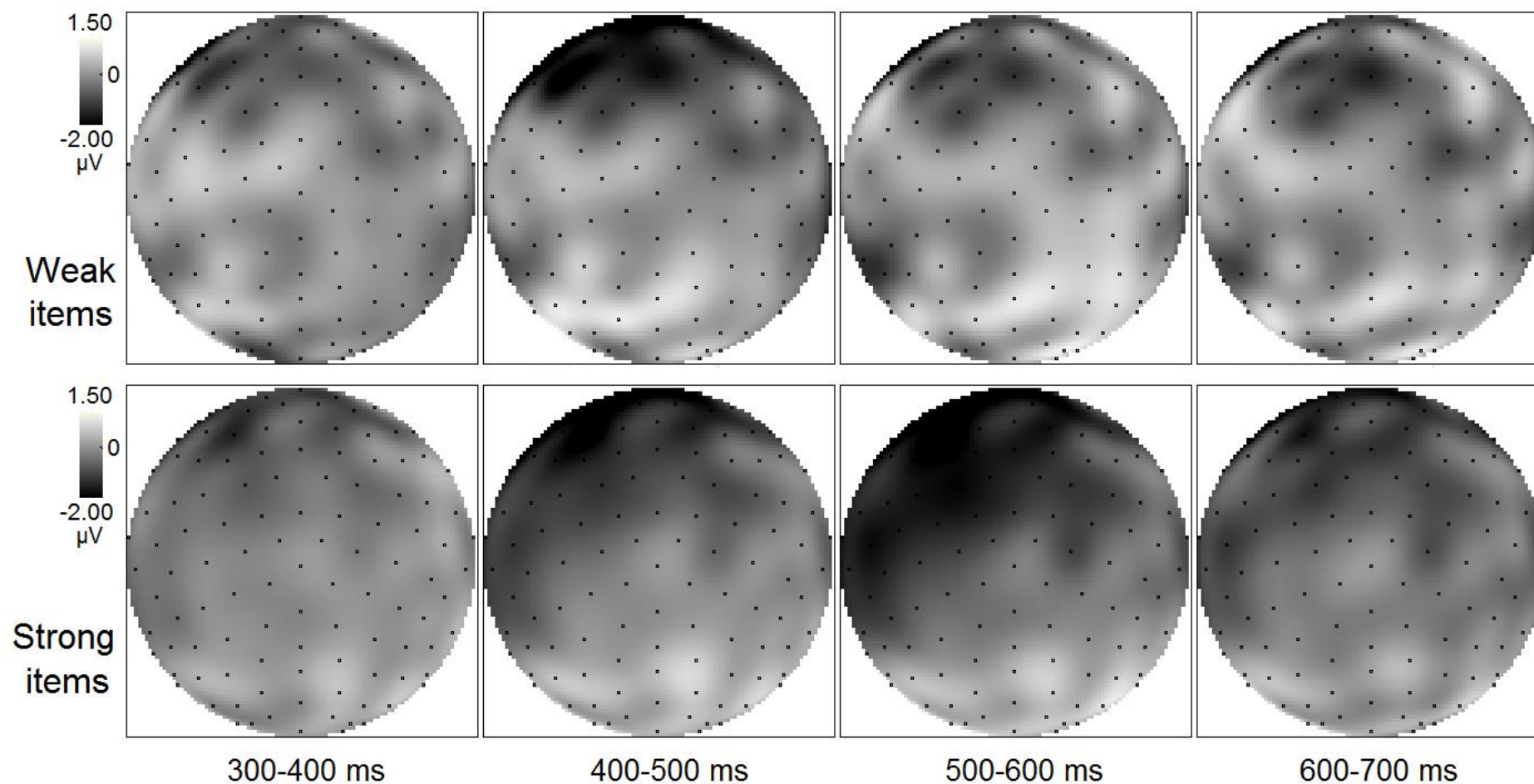


Figure 6-5: Topographical distribution of old/new differences in mean amplitude (μV) for weak items (first row) and strong items (second row) between 300-700ms.

The topography of the old/new effects is illustrated in Figure 6-5, which highlights the posterior distribution of the old/new difference waves elicited by weak and strong items in both the 300-500ms and 500-700ms time windows, and a left hemisphere distribution of the reversed old/new effect for strong items in the 300-500ms time window.

Consistent with previous studies into the neural correlates of familiarity and recollection, and to facilitate comparison with the old/new effects found by other labs (e.g., Allan et al., 1998, Curran, 2000; Friedman, & Johnson, 2000; MacKenzie, & Donaldson, 2007; Mecklinger, 2000; Rugg et al., 1998a; Wilding, & Sharpe, 2003), ERPs were quantified for analysis by computing the mean amplitude relative to the mean of the 200ms pre-stimulus baseline period for 300-500ms and 500-700ms post-stimulus. Separate within-subjects ANOVAs were conducted for the two time windows using electrodes equivalent to F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8 (see Chapter 2, section 2.2.1) in analyses of location (including factors of caudal position, hemisphere, and site). The old/new effects were characterised separately for the weak and strong conditions, and finally comparisons were made between conditions. As ERP effects are only of interest when they reflect differences between conditions, only significant effects involving the factor of condition are reported. The literature reports qualitatively separate old/new effects involving different ERP components in the two time windows; therefore analysis over time was not carried out.

6.2.2.4. *Old/New Effects for Weak and Strong Items*

The old/new effects were characterised separately for the weak and strong conditions. Mean amplitude was analysed in a 2 x 2 x 2 x 3 repeated measures

ANOVA with within-subject factors of item-type (old, new), hemisphere (left, right), site (superior, inferior) and caudal position (frontal, central, parietal).

6.2.2.5. Weak Old vs. New Items

300-500ms Time Window

Analysis of the 300-500ms time window revealed no main effects. There was a significant item-type by caudal position interaction ($F(2,34) = 6.03$, $MSE = 2.02$, $p < .01$, $\eta_p^2 = .262$), reflecting differences between new and weak ERPs at parietal electrodes that were absent at frontal and central sites. *A priori* t-tests showed that these differences were significant at parietal electrodes ($t(17) = 3.74$, $p < .01$, $r = .451$), demonstrating the presence of an early old/new effect at parietal electrodes (see Figure 6-6).

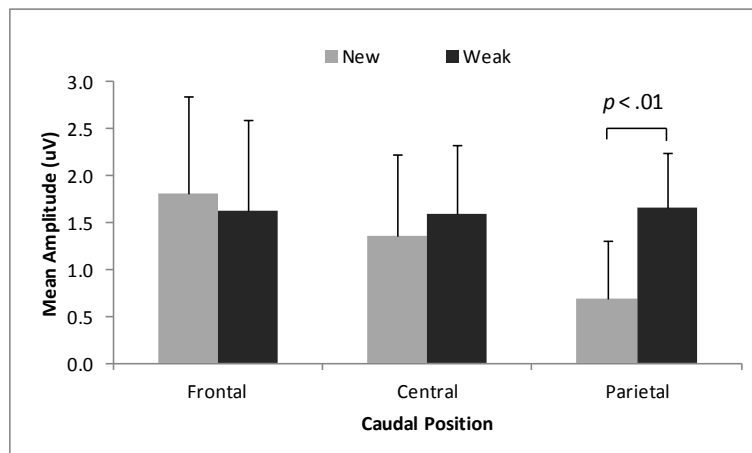


Figure 6-6: Mean amplitude for new and weak items at frontal, central and parietal electrodes at 300-500ms. Error bars show standard error of mean.

The 3-way item-type by caudal position by site interaction just failed to reach significance ($F(2,34) = 2.76$, $MSE = 0.5$, $p = .077$, $\eta_p^2 = .140$) – in general the old/new effect exhibited a superior distribution at parietal sites. Examination of

the data revealed that the old/new effect was maximal at P3 ($t(17) = 2.30$, $p < .05$, $r = .238$).

500-700ms Time Window

Analysis of the 500-700ms time window revealed that the main effect of item-type just failed to reach significance ($F(1,17) = 3.42$, $MSE = 10.2$, $p = .08$, $\eta_p^2 = .167$) – in general the mean amplitude was more positive-going for weak items compared to new items. There was a significant item-type by caudal position interaction ($F(1.35,23.0) = 3.80$, $MSE = 4.23$, $p < .05$, $\eta_p^2 = .183$), reflecting differences between new and weak ERPs at parietal electrodes that were absent at frontal and central sites. *A priori* t-tests showed that these differences were significant at parietal electrodes ($t(17) = 4.07$, $p < .001$, $r = .494$), demonstrating the presence of a late old/new effect at parietal electrodes (see Figure 6-7).

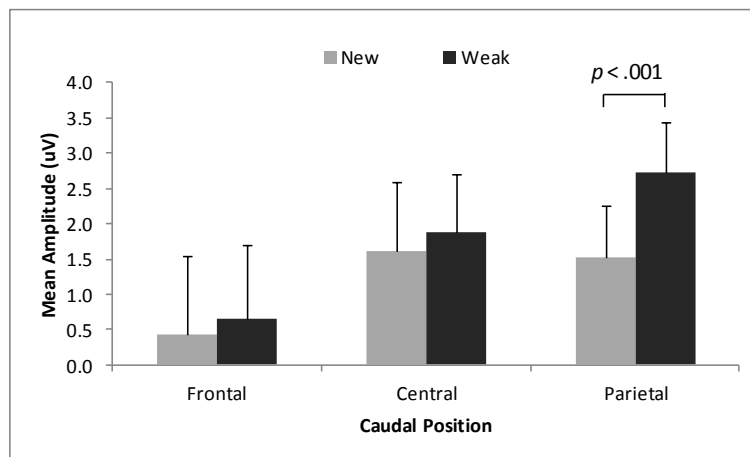


Figure 6-7: Mean amplitude for new and weak items at frontal, central and parietal electrodes at 500-700ms. Error bars show standard error of mean.

The 3-way item-type by caudal position by site interaction was also significant ($F(2,34) = 5.43$, $MSE = 0.917$, $p < .01$, $\eta_p^2 = .242$) – the old/new effect exhibited a superior distribution at parietal sites. Examination of the data revealed that the old/new effect was maximal at P3 ($t(17) = 2.30$, $p < .05$, $r = .238$).

6.2.2.6. Strong Old vs. New Items

300-500ms Time Window

Analysis of the 300-500ms time window revealed a significant item-type by caudal position interaction ($F(1.40, 23.7) = 3.54$, $MSE = 3.75$, $p < .05$, $\eta_p^2 = .172$; Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = 9.08$, $p < .05$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = 0.698$). The interaction reflected larger differences between new and strong ERPs at frontal electrodes than at central and parietal sites. Although on average new items were more negative-going than strong items at parietal electrodes, *a priori* t-tests showed that these differences were not significant ($t(17) = 0.96$, $p > .05$, ns), and new and strong items did not differ significantly at frontal ($t(17) = 1.26$, $p > .05$, ns) or central electrodes ($t(17) = 0.85$, $p > .05$, ns; see Figure 6-8).

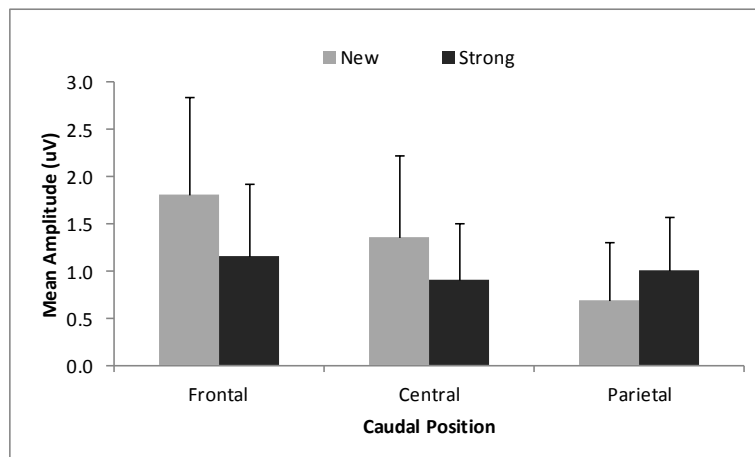


Figure 6-8: Mean amplitude for new and strong items at frontal, central and parietal electrodes at 300-500ms. Error bars show standard error of mean.

The item-type by hemisphere interaction just failed to reach significance ($F(1, 17) = 3.39$, $MSE = 2.39$, $p = .08$, $\eta_p^2 = .166$) – in general the old/new effect was larger over the left hemisphere. The item-type by caudal position by site 3-way interaction was

significant ($F(2,34) = 3.85$, $MSE = 0.631$, $p < .05$, $\eta_p^2 = .185$) – the old/new effect exhibited a superior distribution at frontal sites. Examination of the data revealed that the old/new effect was maximal (albeit reversed in polarity) at F3 ($t(17) = 2.10$, $p < .05$, $r = .206$).

500-700ms Time Window

Analysis of the 500-700ms time window revealed a significant item-type by caudal position interaction ($F(1.47,25.1) = 5.63$, $MSE = 4.12$, $p < .01$, $\eta_p^2 = .249$; Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = 70.7$, $p < .05$), therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = 0.737$)). The interaction reflected larger differences between new and strong ERPs at frontal and central electrodes than at parietal sites. Although on average strong items were more positive-going than new items at parietal electrodes, *a priori* t-tests showed that this difference was not significant ($t(17) = 1.14$, $p > .05$, ns), and new and strong items did not differ significantly at central ($t(17) = 1.53$, $p > .05$, ns) or frontal electrodes ($t(17) = 1.81$, $p = .09$, $r = .162$; see Figure 6-9).

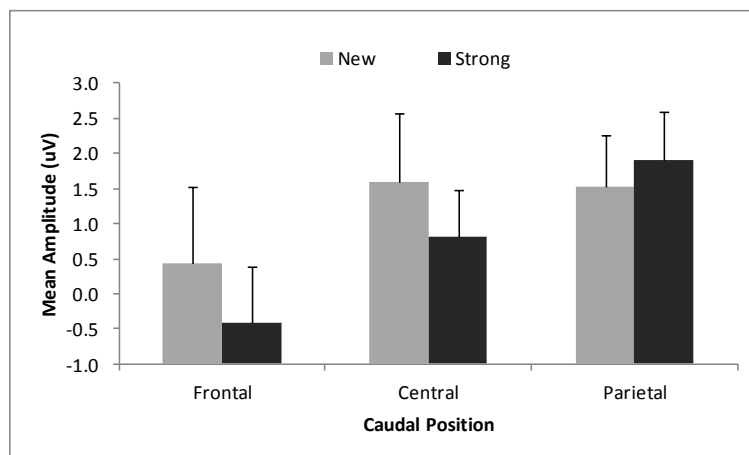


Figure 6-9: Mean amplitude for new and strong items at frontal, central and parietal electrodes at 500-700ms. Error bars show standard error of mean.

The interactions between item-type and hemisphere ($F(1,17) = 3.11$, $MSE = 1.97$, $p = .096$, $\eta_p^2 = .155$), and item-type and site ($F(1,17) = 3.08$, $MSE = 0.827$, $p = .098$, $\eta_p^2 = .153$) just failed to reach significance – in general the old/new effects were larger over the left hemisphere and at superior sites. The item-type by caudal position by site 3-way interaction was significant ($F(2,34) = 6.32$, $MSE = 0.854$, $p < .01$, $\eta_p^2 = .271$) – the old/new effect exhibited a superior distribution at frontal sites.

Examination of the data revealed that the old/new effect was maximal (albeit reversed in polarity) at Fz ($t(17) = 2.65$, $p < .05$, $r = .293$).

6.2.2.7. *Effect of Presentation Frequency on Old/New Effect*

Difference waves were calculated for weak minus new items, and strong minus new items, to allow direct comparison of the magnitude and distribution of the old/new effects in each condition. Mean amplitude difference was analysed in a $2 \times 2 \times 2 \times 3$ repeated measures ANOVA with within-subject factors of condition (weak, strong), hemisphere (left, right), site (superior, inferior) and caudal position (frontal, central, parietal).

300-500ms Time Window

Analysis of the 300-500ms time window revealed that the main effect of condition just failed to reach significance ($F(1,17) = 3.22$, $MSE = 12.0$, $p = .09$, $\eta_p^2 = .159$) – in general the old/new effect was larger in the weak condition than in the strong condition. The interactions between condition and hemisphere ($F(1,17) = 3.91$, $MSE = 2.93$, $p = .065$, $\eta_p^2 = .187$), and condition and site ($F(1,17) = 3.27$, $MSE = 0.61$, $p = .089$, $\eta_p^2 = .161$) just failed to reach significance – in general differences between old/new effects in the two conditions were larger over the left hemisphere and at

superior sites. Examination of the data revealed that the difference between old/new effects in the two conditions was maximal at P3 ($t(17) = 3.05$, $p < .01$, $r = .353$).

500-700ms Time Window

Analysis of the 500-700ms time window revealed a significant main effect of condition ($F(1,17) = 10.5$, $MSE = 9.94$, $p < .01$, $\eta_p^2 = .383$) – the old/new effect was larger in the weak condition than in the strong condition. The 3-way condition by caudal position by site interaction was also significant ($F(2,34) = 4.06$, $MSE = 0.65$, $p < .05$, $\eta_p^2 = .193$), reflecting differences between old/new effects in the two conditions were largest at inferior sites and frontal electrodes. Examination of the data revealed that the difference between old/new effects in the two conditions was maximal at F7 ($t(17) = 2.12$, $p < .05$, $r = .208$).

6.2.2.8. Early Effects (80-150ms)

From visual inspection of the waveforms, a very early negative-going effect was observed for strong items relative to new items at frontal electrodes, between 80-150ms, which resembles the N1 component associated with attention orientation. Therefore an additional analysis was performed within the 80-150ms time window in a 3 x 2 x 2 x 5 repeated measures ANOVA with within-subject factors of item-type (new, weak, strong), hemisphere (left, right), site (superior, inferior) and caudal position (frontal, central, parietal).

Analysis revealed a significant item-type by hemisphere interaction ($F(2,34) = 6.51$, $MSE = 1.796$, $p < .01$, $\eta_p^2 = .277$), reflecting an old/new effect for strong items that was reversed in polarity over the left hemisphere. The 3-way item-type by hemisphere by site interaction was also significant ($F(2,34) = 3.55$, $MSE = 0.235$, $p < .05$, $\eta_p^2 = .173$) – this old/new effect exhibited an inferior distribution for strong items

but a superior distribution for weak items over the right hemisphere. The 3-way item-type by caudal position by site interaction just failed to reach significance ($F(4,68) = 2.17$, $MSE = 0.254$, $p = .08$, $\eta_p^2 = .113$) – in general the old/new difference for strong items exhibited an inferior distribution at parietal sites.

6.2.3. Discussion

Experiment 8 sought to replicate an ERP study which used a memory strength manipulation to produce graded early and late ERP old/new effects, and to extend the study to include concurrent pupil-size recording, with the aim of also observing a graded pupil-size response.

The pupillometry findings showed a main effect of item-type – in general participants' pupils dilated more in response to correctly identified weak old items than to correctly identified new items. Somewhat surprisingly, and counter to predictions, participants' pupils also dilated more to correctly identified weak old items than to correctly identified strong old items, and the difference between new and strong items only approached significance.

ERP findings in the 300-500ms window showed an old/new effect at parietal electrodes for weak items, and although the literature generally reports mid-frontal old/new effects, some studies have reported parietal old/new effects in the 300-500ms time window (MacKenzie, & Donaldson, 2007; Yovel, & Paller, 2004), including the study that was being replicated here (Finnigan et al., 2002). Contrary to predictions, strong items did not show enhanced positivity compared to new items; instead, the old/new effect for strong items, which was exhibited at frontal electrodes, was reversed in polarity – new items were more positive than strong old items.

In the 500-700ms window there was also an old/new effect at parietal electrodes for weak items, consistent with the literature (e.g., Curran 2004; Düzel et al., 1997; Friedman, 2004; Rugg et al., 1998b; Smith, 1993; Trott et al., 1999). This was not, however, the case for strong items which, as for the earlier time window, did not show enhanced positivity compared to new items; instead, the old/new effect for strong items, which was exhibited at frontal electrodes, was reversed in polarity – new items were more positive than strong old items.

To try to understand whether the reversal of the old/new effect for strong items was due to an absence of enhanced positivity of the FN400 and LPC components, or whether it could be due to modification of an earlier component, a follow-up analysis was performed on the 80-150ms window, to see whether the memory-strength manipulation may have influenced early attention processes such as those reflected by the N1. Analysis revealed reversed polarity old/new differences for strong items over the left hemisphere than the right hemisphere. In the later time windows, interactions with hemisphere were not found to be significant, however there was a trend for hemisphere differences in both time windows in the analysis of the triple presentation condition (strong items in relation to new items). In general the strong old/new effect was larger over the left hemisphere than the right hemisphere in both the 300-500ms and 500-700ms time windows.

Comparing the old/new effects between the conditions in the 300-500ms time window, an interaction approaching significance suggested the old/new effect in the weak condition was slightly larger over the left hemisphere than the right hemisphere, whereas in the strong condition the old new effect was much larger over the right hemisphere than the left hemisphere. There appear to be laterality differences between weak and strong words, and single and triple presentation

old/new effects, which begin as early as 80-150ms after stimulus presentation and persist in the later components. The recognition positivity reported in the literature is usually maximal over the left hemisphere, suggesting that in this experiment it may be absent or reduced for strong items.

6.3. Experiment 9 – Pupil and Behavioural Data Only

Experiment 8 yielded the surprising finding that rather than a graded strength of memory effect, both ERPs and pupillometry results demonstrated only a very weak old/new effect for strongly encoded items (those that were repeated three times during study), but a much more robust old/new effect for the weakly encoded items (which were only encountered once during study). Therefore, Experiment 9 was carried out as a near-identical behavioural/pupillometry only replication of Experiment 8, in the laboratory environment used for Experiments 1-7, to determine whether this pattern of results was attributable to the task itself (which differed considerably from those used in previous experiments), or whether it was somehow associated with the ERP testing environment and procedure. For example, the illumination level in the EEG room is low compared to the room in which the other experiments were performed, and it is possible that the procedure itself increased overall levels of arousal in the participants.

6.3.1. Method

6.3.1.1. *Participants*

Thirty-one participants (7 male; age range: 18.2-43.5, $M = 21.7$, $SD = 6.79$) with normal or corrected-to-normal vision were recruited from the psychology course-credit and subject pools at the University of Sussex, and through personal contact.

6.3.1.2. *Materials/Apparatus/Design/Procedure*

Stimuli, design and procedure were the same as in Experiment 8 (see sections 6.2.1.2 and 6.2.1.3) with the following changes: Study and recognition lists were presented in black 20pt Monospaced font on a light grey background under fixed illumination. Words were presented using Experiment Builder software (SR-Research, Ontario) on a 21" CRT monitor. Participants viewed the monitor from a distance of 70cm and the visual angle subtended by the words was approximately 3°. Eye movements were recorded with an EyeLink II (SR-Research, Ontario), with a sampling rate of 500Hz. All items are presented in Appendix G.

6.3.1.3. *Pupil Recording*

Maximum pupil-size was recorded from the right eye during each recognition period. A Pupil Dilation Ratio (PDR; see Chapter 2, section 2.1.2.1) was calculated expressing the maximum pupil-size for each 2000ms recognition trial as a proportion of the maximum pupil-size during that trial's 200ms baseline.

6.3.2. **Results**

6.3.2.1. *Behavioural Data*

The proportions of correct responses to new, weak (presented once at learning) and strong (presented three times at learning) items were calculated and averaged 80.4% ($SD = 8.3\%$) correct new items, 54.1% ($SD = 13.3\%$) correct weak items and 73.9% ($SD = 13.5\%$) correct strong items. A one-way repeated-measures ANOVA revealed a significant main effect of item-type ($F(1.14, 34.3) = 53.1$, $MSE = 0.019$, $p < .001$, $\eta_p^2 = .639$). Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(2) = 40.3$, $p < .001$), therefore degrees of freedom were corrected

using Greenhouse-Geisser estimates of sphericity ($\epsilon = 0.571$). Bonferroni-corrected subsidiary t-tests revealed that in general participants correctly identified significantly fewer weak items than strong ($t(30) = 20.36$, $p < .001$, $r = .932$) or new items ($t(30) = 8.29$, $p < .001$, $r = .696$; see Figure 6-10).

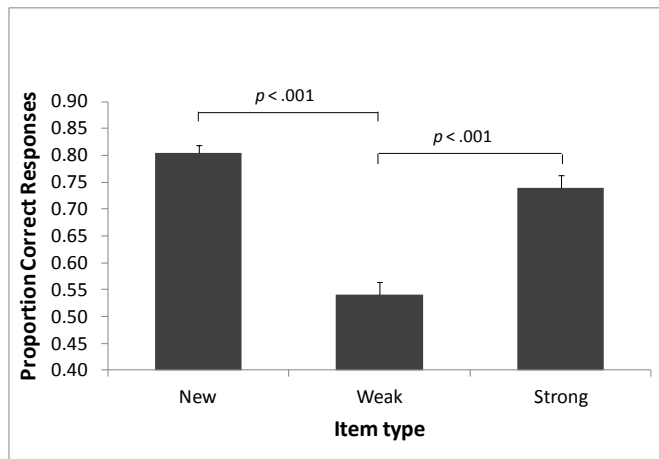


Figure 6-10: Proportion of correct responses to new, weak and strong items. Error bars show standard error of mean.

6.3.2.2. Pupil-Size Data

Average PDR for old and new items was calculated for correctly identified new, weak and strong items. As PDR is a function of baseline pupil-size, baseline pupil-sizes for old and new items in each condition were compared to ensure that any differences in PDR were not due to baseline differences. The difference was not significant ($F(2,60) = 1.69$, $p > .05$, ns, $\eta_p^2 = .053$).

Average PDR for correctly identified new, weak and strong items were compared in a one-way repeated measures ANOVA, which showed a main effect of item-type ($F(2,60) = 7.07$, $MSE < 0.001$, $p < .01$, $\eta_p^2 = .191$). Planned t-tests revealed that in general participants' pupils dilated more to correctly identified strong old items ($M = 1.076$, $SD = 0.0278$) than to correctly identified new items ($M = 1.064$, $SD = 0.0306$,

$t(30) = 4.85, p < .001, r = .440$). There was a trend for participants' pupils to dilate more to correctly identified strong old items than to correctly identified weak items ($M = 1.069, SD = 0.0292; t(30) = 1.90, p = .06, r = .108$; see Figure 6-11).

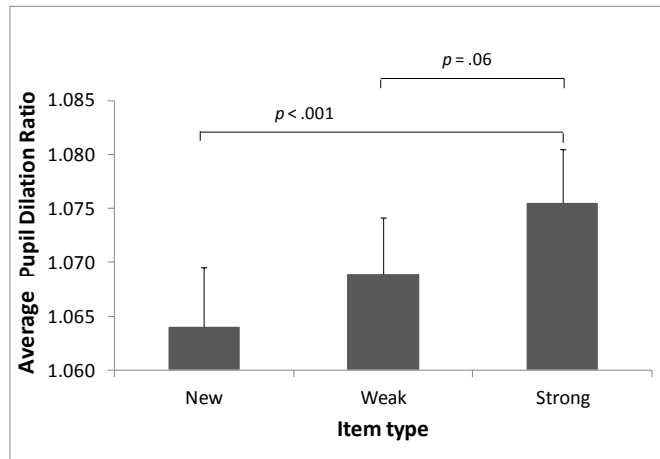


Figure 6-11: Pupil dilation ratio for correctly identified new, weak and strong items. Error bars show standard error of mean.

6.3.3. Discussion

Experiment 9 was designed to establish whether the strength of memory manipulation adopted in Experiment 8 (Finnigan et al., 2002) would produce a graded pupil response in addition to the old/new effect between weak and new items. As in Experiment 8, in general participants correctly identified more strong and weak items than new items, and the proportion of correctly identified items in each category was roughly the same (new: 81.2% vs. 80.4%; weak: 57.2% vs. 54.1%; strong: 72.9% vs. 73.9%).

Participants' pupils dilated more to correctly identified strong old items than to correctly identified new items. There was also a trend for participants' pupils to dilate more to correctly identified strong old items than to correctly identified weak items, with no significant difference between weak and new items. This result is in contrast to Experiment 8 where participants' pupils dilated more to correctly

identified weak old items than to correctly identified strong old items and correctly identified new items. As such, it suggests that some feature of the ERP testing environment led to the unexpected finding of a stronger pupil (and ERP) response to the weak compared to strong items in experiment 8. This possibility is explored further in the general discussion.

6.3.4. General Discussion

Experiment 8 sought to replicate an ERP study which used a memory strength manipulation to produce graded early and late ERP old/new effects (Finnigan et al., 2002). Using a novel approach, it also sought to extend the study to include concurrent pupil-size recording, with the aim of also observing a graded pupil-size response, which might suggest that the two psychophysiological measures are linked. Concurrent measurement was designed to allow the memory processes from the same group of participants, during precisely the same task, to be quantified in two different psychophysiological indices of recognition memory.

Results showed that for weak items (studied once during learning) there was a PONE, and early and late ERP old/new effects at parietal electrodes. However, contrary to predictions, for strong items (studied three times during learning) early and late ERP old/new effects occurred only at frontal electrodes and were reversed in polarity; in addition the PONE for strong items was absent. Therefore the predicted graded strength of memory effect was not obtained in either ERP or pupillometry data. In order to test the effects of the strength of memory manipulation on pupil-size alone, Experiment 9 was carried out as a near-identical behavioural/pupillometry only replication of Experiment 8, in the laboratory environment used for Experiments 1-7. Experiment 9 showed that in contrast to Experiment 8 there was a PONE for strong old items – participants' pupils dilated

more to correctly identified strong old items than to correctly identified new items. There was also a trend for participants' pupils to dilate more to correctly identified strong old items than to correctly identified weak old items, with no significant difference between weak and new items.

The strength of memory manipulation was expected to be effective because stimulus repetition has been shown to enhance memory performance on behavioural measures (e.g., Leding, & Lampinen, 2009; Yonelinas, 2002) and previous studies (Finnigan et al., 2002; MacKenzie, & Donaldson, 2007; Van Petten et al., 1991) found graded ERP components using memory strength manipulations. However it did not appear to work in Experiment 8. Analysis of the ERP mean amplitudes for weak and new items revealed a parietal old/new effect at 300-500ms, and although the literature generally reports frontal/central old/new effects, some studies have reported parietal old/new effects in the 300-500ms time window (MacKenzie, & Donaldson, 2007; Paller, Gonsalves, Grabowecky, Bozic, & Yamada, 2000; Yovel, & Paller, 2004), including the study that was being replicated here (Finnigan et al., 2002). The 500-700ms window demonstrated a parietal LPC old/new effect for weak items, consistent with the literature (e.g., Curran, 2004; Düzel et al., 1997; Friedman, 2004; Rugg et al., 1998b; Smith, 1993; Trott et al., 1999).

Analysis of the ERP mean amplitudes for strong and new items showed that strong items did not show the expected enhanced positivity compared to new items in either the 300-500ms or 500-700ms time window. Instead new items showed enhanced positivity compared to strong items at frontal electrodes, reversing the polarity of the old/new effect. This meant that the predicted graded memory effect did not occur for ERPs at frontal, central or parietal electrodes. This was echoed in the pupil-size data which showed a significant old/new difference for weak but not

strong items, similar to ERPs at parietal electrodes. Norman, Tepe, Nyhus and Curran (2008) strengthened memory for some stimuli by presenting faces three times (weak) or six times (strong) during learning. They demonstrated that the strength manipulation had no impact on the FN400 old/new effect (300-500ms), and that the LPC old/new ERP effect (400-800ms) was only present for weak items – the positive-going waveform evoked for strong items was equivalent in magnitude to the waveform for new items. Because participant-reported “remember” judgements were reduced by the manipulation, whilst “know” judgements were not, Norman et al. (2008) concluded that the interference caused by increasing the “strength” of items affected recollection but not familiarity, and reduced the LPC old/new effect. This explanation is consistent with the absence of a parietal LPC old/new for strong items in Experiment 8, however it does not explain why the expected FN400 old/new effect was also missing for strong items, or why in both time windows strong old/new effects were reversed at frontal electrodes.

Although quantitatively different, weak and strong items demonstrated qualitatively the same topographical changes, whereas the distribution for new items was qualitatively different to both strong and weak items. This could indicate that an earlier latent ERP component was modified for strong items as a result of the triple presentation (for example an early negative component being larger, or an early positive component being smaller than for weak items). This was explored in a follow-up analysis in the 80-150ms time window (selected by inspection of the grand-averages), to see whether the memory-strength manipulation may have influenced early attentional processes such as those reflected by N1. Analysis of the 80-150ms window revealed early old/new differences over the left hemisphere for strong but not weak items. In the FN400 and LPC time windows (300-500ms and 500-700ms) hemisphere by condition interactions also showed a trend towards

larger old/new differences over the left hemisphere for strong items. There appear to be laterality differences between weak and strong words, and single and triple presentation old/new effects, which begin as early as 80-150ms after stimulus presentation and persist in the later components FN400 and LPC.

If stimulus repetition meant that fewer processing resources needed to be allocated to strong items, then this might account for differences between weak and strong items in the 80-150ms time window. Target stimuli can be sorted and selected for additional processing at a very early stage, so early components such as the N1 are affected by attention (Hillyard, Hink, Schwent, & Picton, 1973; Näätänen, 1990), therefore if a repeated strong stimulus becomes less interesting, and therefore less attended, the N1 might be attenuated.

Repetition may lead to priming, whereby subsequent presentations of an item are processed more quickly and efficiently (cf., implicit memory, Chapter 3; Tulving, & Schacter, 1990), however, negative priming can lead to poorer episodic encoding for highly primed items (Tipper, 1985; Wagner, Maril, & Schacter, 2000). A related phenomenon is repetition suppression, where repeated stimuli produce less neural activation than new stimuli (for review, see Schacter, & Buckner, 1998; Buckner et al., 1998; Grill-Spector, & Malach, 2001; van Turennout et al., 2000) or stimuli presented once at learning (e.g., weak; Jiang, Haxby, Martin, Ungerleider, & Parasuraman, 2000).

Monkey single-cell recordings (Desimone 1996; Miller, & Desimone, 1994) and human fMRI studies have shown that repetition leads to decreased activation in brain regions involved in stimulus processing, such as the left inferior prefrontal cortex (LIPC; Wagner, Desmond, Demb, Glover, & Gabrieli, 1997) and hippocampal

and parahippocampal regions (Brozinsky, Yonelinas, Kroll, & Ranganath, 2005; Suzuki, Johnson, & Rugg, 2011), and is suggested to occur automatically during learning (Wiggs, & Martin, 1998). Guo, Lawson and Jiang (2007) found that the late posterior ERP repetition effect (>550ms) showed more positive-going ERP amplitudes to items at initial presentation compared to repetition.

Whilst the pupil response was not graded as expected, it did echo the ERPs at parietal electrodes in that there was an old/new effect in the expected direction for weak items, and weak items elicited a larger response than strong items. The similar pattern of results in the ERPs and PDR old/new effects suggests that both measures index the same underlying cognitive processes, and, importantly, lends support to Otero et al.'s (2011) argument that the PONE represents neurocognitive activity underlying recognition memory. Similarly, like the potential ERP repetition suppression effect which may explain findings in Experiment 8, Van Rijn, Dalenberg, Borst, and Sprenger (submitted) found that the phasic pupil response to repeated stimuli decreased by 2% per repetition, which could account for the equivalent effect seen in the pupil response.

Behavioural performance was consistent between the two experiments, suggesting that the implementation of Experiment 9 in a different laboratory was similar enough to produce comparable task performance, and that the ERP procedure and equipment did not massively distract participants from the stimuli. However, the manipulation did not produce the expected graded pupil-size in either experiment, although the results of Experiment 9 were more in line with the expected effect with a PONE for strong items and weak items on average being intermediate in size between strong and new (albeit non-significant).

A possible influencing factor in Experiment 8 was the use of a remote eye-tracker. Unlike the head-mounted eye-tracker used in Experiment 9, remote cameras are usually located 50-100cm from the eye, and therefore measure the pupil with lower precision than head-mounted eye-trackers (see Chapter 2, section 2.1.1.1; Klingner, 2010; Marshall, 2002). It is possible that the use of a remote eye-tracker in these experiments (necessitated by the ERP acquisition) reduced the precision of the pupil measurement, and perhaps underestimating pupil-size when it was at its largest (e.g., for strong items). Eye-tracking during Experiment 8 was more vulnerable to loss of signal due to wires across the participant's face and change of focus, and was less amenable to correction except for between recording blocks for fear of loss of ERP data. There was a 2% loss of correctly eye-tracked new trials only (81.2% correct trials vs. 78.7% eye-tracked correct trials; $t(17) = 10.7$, $p < .001$, $r = .871$), and the total number of tracked strong trials was 98.9%, compared to 99.8% weak ($t(17) = 4.50$, $p < .001$, $r = .544$) and 100% of new trials ($t(17) = 3.43$, $p = .003$, $r = .409$). Additionally, the ambient illumination in the room used for Experiment 8 was lower than that used for Experiment 9, which would have influenced pupil-size.

Although the number of artefact-free correct trials as a percentage of correct trials per item-type (and therefore the number that were included in each grand average) was not significantly different (all t s < 2 , ns), there were significantly fewer absolute numbers of trials contributing to the ERP grand averages for weak items ($M = 51.9\%$, $SD = 15.1\%$) than for new ($M = 70.7\%$, $SD = 15.4\%$; $t(17) = 3.91$, $p = .001$, $r = .474$), or strong ($M = 64.3\%$, $SD = 14.3\%$; $t(17) = 8.53$, $p < .001$, $r = .810$) items. Although issues may arise with waveforms formed from differing numbers of trials, because of the resultant differing signal-to-noise ratios (Luck, 2010; see Chapter 2, section 2.2.2), this is more of a concern when measuring peak amplitude due to the greater influence a spurious peak has over the peak measurement in averages

containing fewer contributing trials (Luck, 2010). Experiment 8 measured mean amplitude, an unbiased measure even when trial numbers differ, and means that this difference in numbers is unlikely to have biased the results (Luck, 2010).

In considering why the results of Experiment 8 are different to those of the replicated study by Finnigan et al. (2002) several factors should be acknowledged. Although as far as possible the study design and procedure were replicated, ultimately there were methodological differences, for example Finnigan et al. (2002) used a 30 electrode cap with lower maximum impedance than the 128 electrode net used here. Presentation duration was increased from 400ms to 1000ms at both learning and recognition to bring the procedure in line with Experiments 1-7, and a mask of “HHHHH” preceded and followed stimuli at recognition, rather than the blank screen used by Finnigan et al. (2002), in order to minimise the influence of the light reflex. Although they were the same length (5 letters), different stimuli were used in Experiment 8 to those used by Finnigan et al. (2002), and font size and distance from screen may also have differed as these were not provided. Therefore it is possible that basic visual stimuli features influenced the ERP waveform, particularly with respect to early visual components (Luck, 2005; Schloerscheidt, & Rugg, 2004).

Experiment 8 used an online vertex reference electrode, whereas Finnigan et al. (2002) used physically linked earlobe electrodes as an online reference, which although not biased towards either hemisphere, creates “a zero-resistance electrical bridge between the hemispheres, distorting the voltage distribution and reducing hemispheric asymmetries” (Luck, 2005, p. 107). Electrodes in Experiment 8 were re-referenced offline to a virtual average mastoid after the left and right mastoid recordings were checked for artifacts, whereas Finnigan et al. (2002) do not appear to have re-referenced offline.

In the separate analyses of their two experiments, Finnigan et al. (2002) collapsed data across responses, whereas the data for Experiment 8 was analysed for correct items only to limit analysis to items most likely to elicit a genuine memory. In their analysis of correct and incorrect trials, Finnigan et al. (2002) collapsed data over their two experiments because they had insufficient incorrect trials to form a grand average in either experiment separately. Although their experiments were very similar (procedurally only the length of the recognition list varied), the study-test repetition lag varied between experiments, and lag can itself influence repetition effects such as suppression (e.g., Brozinsky et al., 2005).

Experiment 8 and the experiments by Finnigan et al. (2002) took place in different laboratory environments with different researchers, equipment and sources of noise, and ultimately a different group of participants, therefore it would not necessarily be expected that the two produce the same results for any or all of the above reasons. In addition, contrary to Finnigan et al. (2002), Opitz (2010) found no difference in late parietal old/new ERP effects between items presented once and items presented three times.

Despite focusing on two components, FN400 and LPC, repetition of learning items in this manipulation may have had wider influence than the single 'memory strength' effect intended. If the manipulation affected more than these two components, then this could explain the apparent lack of difference between new and strong items, for example modulation of overlapping positive- or negative-going ERPs. The issue of latent components makes it more difficult to interpret the waveform, which is a local sum of voltage differences. A reduction of the amplitude of the FN400 or LPC components, as manifest in the grand-average, may not reflect a reduction in the underlying neural activity of interest (see Luck, 2005, for further discussion).

Although the literature reports separate ERP components in the 300-500ms and 500-700ms time windows analysed, very similar patterns of results were reported for the two time periods in Experiment 8. An interesting development of Experiment 8 might be to perform an analysis over time, to test whether or not the early and late old/new effects were statistically different, therefore reflecting separate components.

7. General Discussion

Conclusions, Limitations and Priorities

7.1.1. Summary

The central aim of this thesis was to explore the cognitive processes associated with the recently reported Pupil Old/New Effect (PONE), whereby the pupil dilates to a larger maximum size in relation to a baseline when participants view old items compared to when they view new items during a recognition memory test. This concluding chapter will summarise the key results of Experiments 1 to 9, noting some of the limitations, and relating the findings back to the main issues in pupillometry and recognition memory research outlined in the introductory chapter. Finally, it will offer some suggestions for the future direction of this research.

Experiments 1 and 2 set out to replicate the PONE observed in explicit tests of recognition memory and determine whether it would also be present in an “implicit” test of memory using perceptual fluency. Results showed that the PONE was replicated in a standard test of recognition memory, but not in an “implicit” test of perceptual fluency, and it did not occur when participants were asked to read word stimuli rather than make a recognition decision. Experiments 3 and 4 extended this finding by examining whether the PONE would be present when recognition memory was tested using artificial grammar learning, a form of implicit learning that relies on a sense of familiarity to facilitate recognition. The PONE was again replicated in a standard explicit recognition task, but was not present when participants were judging grammatical vs. ungrammatical letter strings. Experiments 5 and 6 examined the effects of asking participants to deliberately perform poorly during recognition, and crucially demonstrated that the PONE is still present when

participants are asked to give false behavioural answers in a malingering task, and even when they are asked not to respond at all, but is absent if participants randomly preload an answer without making a recognition decision. A further important finding from Experiment 7 was that despite a slight impairment of performance levels, the PONE is still present when attention is divided both at learning and/or during the recognition phase. Experiments 8 and 9 set out to explore whether the PONE and concurrent ERPs responded in a graded manner to an ERP memory strength manipulation, and showed that the PONE is accompanied by parietal ERP old/new effects at 300-500ms and 500-700ms, showing enhanced positivity for old items presented once at learning, compared to new items, and that neither the PONE nor the ERP old/new effects are enhanced by repetition of items during learning.

Across all the experiments reported in this thesis that employed a standard recognition memory procedure, maximum pupil-size was larger when participants looked at old items compared to when they looked at novel items. Taken as a whole, these findings support the theory that the PONE reflects mnemonic processes recruited when participants make a recognition decision. It is important to note that even when an item is new, mnemonic processes are activated – in part due to prior exposure to the common English words used in the experiments, but also because participants are actively seeking to reject novel items, for example searching their memory to ensure that the item was **not** presented. This may account for the fact that pupil-size to new items, in conditions where participants make a recognition decision, was still often larger than pupil size to old or new items in conditions where no recognition decision was required (such as reading, although in the short duration reading condition of Experiment 2 pupil size was larger than for

the short duration recognition condition). In our results we have evidence for both effort effects and mnemonic effects.

There is an extensive literature documenting the effects of “cognitive effort” on pupil size (see Chapter 1, section 1.2.2.1) with some authors proposing that the PONE is nothing more than the result of the greater cognitive effort required to correctly identify old compared to new stimuli. Võ et al. (2008) suggest that recollection requires the retrieval of qualitative contextual information, including the experience of an old item during the study phase, which is more cognitively demanding than the correct rejection of a new item, which does not. Their theory would predict that the PONE should be smaller for deeply encoded items than shallowly encoded items because less effort is required for recollection, however this was not what was found when tested in Experiment 1. The central argument of this thesis, therefore, is that the PONE is the result of conscious recollective processes that accompany the recognition decision, and that items that are better remembered, or have a “stronger” memory, are associated with a larger pupil-size (in line with Otero et al., 2011; Papesh et al., 2011).

Although the experiments reported here did not directly measure participants’ introspective remember-know judgements, Experiments 3 and 4 used artificial grammar, which Scott and Dienes (2008) propose elicits decisions based on familiarity in the absence of recollection. In these experiments, no PONE was found in response to familiar versus unfamiliar grammatical strings. This finding suggests that within a dual-process model of recognition memory, the PONE reflects primarily recollective processes. Whilst others (e.g., Otero et al., 2011) have found a larger pupil size in response to old items rated as “known”, compared to new items, at trend levels, it is difficult to exclude the possibility of recollective experience

contaminating familiarity judgements, with weaker memory strength leading to both the intermediate pupil size and the failure to say that the item is remembered.

Wixted and Mickes (2010) propose that the R-K paradigm merely distinguishes strong and weak memories. It is therefore possible to interpret the findings of this thesis in line with continuous strength models of recognition memory, such as the signal detection unequal variance model (Wixted, 2007a), STREAK (Rotello et al., 2004), and single-trace dual-process models (e.g., Greve, Donaldson, & van Rossum, 2010). These models assume that like familiarity, recollection also lies on a continuum, and that rather than recognition decisions being based on either recollection OR familiarity, both sources of memory information are summed into a unitary combined memory strength that is then compared with a criterion value to make a recognition decision (Wixted, & Stretch, 2004; Wixted, 2007a).

Experiment 8 attempted to provide further evidence for a memory strength signal in the pupil by replicating a graded memory strength ERP study, which demonstrated greater positivity for strongly encoded items relative to items with weaker encoding (Finnigan et al., 2002), and concurrently measuring a graded pupil response. For weak items an enhanced positivity relative to new items was present at parietal electrodes and maximal at P7 in both time windows. Whilst this is consistent with the left parietal old/new effect seen in the literature around 500-700ms, the 300-500ms old/new effect for word recognition is typically seen maximally at fronto-central sites rather than at parietal sites (e.g., Curran, 2000). Other recent studies have also shown an early old/new effect with a posterior scalp distribution, however these studies have been concerned with face recognition rather than word recognition (MacKenzie, & Donaldson, 2007; Paller et al., 2000; Yovel, & Paller, 2004). Finnigan et al. (2002) found a posterior old/new effect between 300-500ms for word stimuli, but did not discuss possible origins, they merely referred to it as a

“posterior N400 strength effect” (p. 2300) and interpreted their results within a single-process model of recognition memory. Although a posterior distribution for the early old/new effect might suggest that the same cognitive processes underlie familiarity and recollection, as MacKenzie and Donaldson (2007) point out, even when the topography of ERPs overlap in this way, it is not possible to determine whether or not common neural generators are implicated.

Although the expected old/new effects were not present for strong items, early and late frontal old/new effects of reversed polarity were demonstrated. In addition, very early hemisphere differences (80-150ms) between weak and strong items persisted throughout the trial, reflecting a difference in magnitude and/or location of neural activity. The different pattern of old/new effects for strong items may be due to interference effects, as proposed by Norman et al. (2008), leading to a reduction in recollection. An alternative explanation is that one or both old/new effects has been attenuated by repetition suppression, where less neural activity occurs for repeated stimuli (e.g., strong) compared to new stimuli (see Schacter, & Buckner, 1998) or stimuli presented once at learning (e.g., weak) (Jiang et al., 2000). Repetition suppression is well documented in the ERP and other psychophysiological literatures (e.g., Guo et al., 2007; Suzuki et al., 2011), and has recently been reported in the pupil literature (Van Rijn et al., submitted).

Although the graded memory manipulation didn't produce the predicted pattern in either the ERP data or pupil-size data, the two psychophysiological measures did respond in qualitatively the same manner, producing old/new differences for weak items but not strong items at parietal electrodes. The parallel occurrence of an old/new effect for weak items, and a possible repetition suppression effect for strong items, in both the ERP and pupil-size data, raises the possibility that the two

measures index the same underlying neurocognitive processes occurring during recognition memory.

Experiments 1 and 5 included measures of participant confidence, which is a potential covariant to pupil size, through the positive emotion associated with a correct response (Kahneman, 1973; Muldner et al., 2009). As discussed in Chapter 1, section 1.2.3.1 and Chapter 3, section 3.2.3, another link between confidence and pupil-size is memory strength. If, as argued earlier, pupil-size during explicit recognition decisions reflects the strength of the underlying memory trace, it is to be expected that trials that have a “strong” memory lead both to a high level of confidence and a larger PDR. In other words, if confidence ratings are taken as a reflection of participants’ subjective experience of the strength of this aggregate signal, and the pupil-size increase reflects the cognitive processes that drive this signal, then pupil-size increases should be greater for high compared to low confidence judgments, as was indeed the case. Interestingly, when analyses were restricted to highly confident answers only, pupil size was still significantly larger for old items than new items, suggesting that confidence and pupil-size may both be downstream effects of memory strength but remain, to some extent at least, independent.

Beatty and Wagoner (1975; 1976) measured confidence and pupil-size in a target detection task and found that largest pupil-sizes were evoked by highly confident hits and the smallest pupil-sizes occurred for highly confident rejections, with low confidence hits and misses in between, suggesting that confidence and pupil-size are not always tightly coupled. In addition, the Quiet condition of Experiment 6 demonstrated a PONE without a behavioural response, suggesting that the PONE cannot simply be the result of confidence in a correct response. An experiment that

could potentially explore this idea further is the Deese-Roediger-McDermott (DRM) paradigm (Roediger, & McDermott, 1995), which elevates the rate of false alarms (new items identified as old), and therefore confidence in items which are new but are confidently identified as old. Using this paradigm, Otero et al. (2011) found that average PDR for false alarms was significantly smaller than for correctly identified old items, therefore a replication which also measured confidence might show that false-alarms have a significantly smaller average pupil-size but an equivalently high confidence rating to old items.

There are a number of methodological issues within these that, with the benefit of hindsight, could have been improved. As acknowledged in Chapter 3 section 3.2.3, the larger pupil size to new items than old items in the Implicit condition of Experiment 1 may have been a novelty or orienting response, visible in the absence of the PONE (Laeng et al., 2007; Lynn, 1966; Pavlov, 1927; Sokolov, 1963). However, as this finding was not replicated in the short duration reading condition of Experiment 2, it is difficult to draw conclusions about the underlying cause without further replication. Given the effect sizes seen in the Implicit condition of Experiment 1, Experiment 2 had sufficient participants ($n = 28$) to give a statistical power to detect similar sized effects of 80%.

The pupil results of Experiments 8 and 9 are slightly contradictory in that for Experiment 8 the PONE was only present for weak items, but in Experiment 9 the PONE was only present for strong items. In addition, in Experiment 8, pupil size for weak items was also significantly larger than for strong items, however this was echoed by the ERP data which also showed an old/new effect for weak items only and more positive-going ERPs for weak than strong items. PDRs in Experiment 9 were smaller than expected, and given the effect sizes seen in Experiment 8,

Experiment 9 needed a larger number of participants ($n = 56$) to increase the statistical power to 80% to detect the same size effect.

It should be acknowledged that different outcomes can arise from different analyses of the same pupil-size data. One limitation of the analyses included in this thesis is that they focussed on the main effect of item-type, comparing PDR to old items with PDR to new items. Alternative interpretations and conclusions may have been drawn had data been analysed and presented in terms of interaction effects, i.e. differences in PDR *between* old and new items in each condition. This approach is equivalent to the construction and comparison of difference waves to analyse ERP data, such as in section 6.2.2.7 of Chapter 6, and would be carried out by subtracting PDR to new items from PDR to old items and subjecting the PONE subtraction data to the ANOVA. For example, the results of Experiment 1, summarised in Figure 3-4, highlight the difference in overall PDR to correctly identified new items, which in the Implicit condition is larger than in the Explicit condition. This has the effect of detracting from the old/new effect, and provides support for the effort account of the PONE (Võ et al., 2008). Had the data been presented as old/new differences, then it would more clearly show support for the strength account of the PONE (Otero et al., 2011) because the PONE in the Explicit condition is larger than in the Implicit condition.

In addition, had the analysis of the Levels Of Processing (LOP) manipulation, illustrated by Figure 3-5, been presented as old/new differences, it would have been clear that the PONE was larger for deeply encoded items than shallowly encoded items in the Explicit condition – better supporting the argument that deeply encoded items are associated with a stronger memory signal. In Experiment 2, the results analysed and presented in Figure 3-10 appear to emphasise an effort-related main

effect of short vs. long presentation duration, with short duration conditions requiring more cognitive effort. Had data been presented as old/new differences in each condition, then it would have been clear that the difference was largest for the two recognition conditions, with the strongest memory signal for the long duration recognition condition. Similarly stronger support for the memory strength account of the PONE can be made by alternative presentations of the data for the remaining seven experiments.

In light of this issue, it is important to reflect on the balance of evidence, and how the data presented here speaks to the effort vs. memory strength debate outlined in the introductory chapter. As discussed above, evidence for the effort account of the PONE can be found in the results of Experiment 2, which show a larger overall pupil-size to items presented in the more difficult short duration conditions than in the easier long duration conditions. In addition, the pattern of results in Experiment 8 could reflect effort-related changes in pupil-size. Weak items were the hardest to correctly identify, as demonstrated by the behavioural data (57.2% vs. 81.2% for new and 72.9% for strong items), and mirroring the changes in pupil-size which were largest for weak items and not significantly different for new and strong items. However, behavioural performance in Experiment 9 was very similar to that of Experiment 8 (54.1% weak, 80.4% new and 73.9% strong items), yet the pupil-size data showed a different result – largest for strong items and not significantly different for new and weak items. In addition, in Experiment 8 pupil responses exhibited the same pattern as the late ERP old/new effect at parietal electrodes, associated with recognition memory processes in an extensive literature (e.g., Curran, 2004; Düzel et al., 1997; Friedman, 2004; Rugg et al., 1998b; Smith, 1993; Trott et al., 1999). If behavioural performance is to be taken to indicate task difficulty, on some level at least, then additional evidence in support of a memory strength account of the

PONE comes from Experiments 3 and 4. Here, performance in the Implicit condition was much lower than that in the standard explicit recognition task (only 61.9% compared to 74.6%), yet pupil-size increase was larger in the Explicit condition. A similar line of argument can be used in interpreting the findings of Experiment 7, in which the divided attention manipulation reduced behavioural performance but did not affect pupil-size between conditions.

In Experiment 5 pupil-size is largest in the Standard condition, despite the fact that in the Malingering condition participants are carrying out a more complex task, involving suppressing a number of correct responses and implementing a covert malingering strategy. Deception itself has been associated with increased cognitive effort (e.g., Dionisio et al., 2011) and/or increased anxiety (e.g., Berrien, & Huntington, 1943), therefore it seems unlikely that the effort explanation of the PONE can account for the largest pupil-sizes occurring in the Standard condition. Further evidence comes from the Incomplete Effort condition of Experiment 6, where participants were required to not pay attention to stimuli but only during the learning phase – at recognition they were to try their best to correctly identify the items. Therefore, the smaller PDR that occurred in the Incomplete Effort condition compared to the Standard condition, is unlikely to be the result of reduced effort and instead fits with the memory strength account of the PONE. The experimental manipulation that was introduced specifically to provide evidence for one account over the other was varying the LOP in Experiment 1. The effort account of the PONE predicts that pupil-size for deeply encoded items should be smaller than for shallowly encoded items because they are easier to remember. In contrast, the results showed that pupil-size for deeply encoded items was larger than for shallowly encoded items, supporting the memory strength account of the PONE.

PDR increased in response to all tasks, in all experiments across the thesis. As discussed in Chapter 1, a vast number of factors influence pupil-size, many of which cannot be controlled or easily isolated in an experimental context – including increases associated with: making an overt response, participant confidence or anxiety, variations in attention, and cognitive effort related to either task difficulty or how hard the participant is trying on a particular trial. It is not the intention to argue that all task-evoked increase in pupil-size is due to memory strength; rather that the balance of evidence suggests that in these experiments the increase in participants' pupil-size when they correctly identify old items compared to new items is predominantly due to a memory-strength signal (e.g., Otero et al., 2011). With hindsight the subtractive difference analysis would have been a stronger way of presenting this central argument.

7.1.2. Future Directions

Future research could extend the present findings in several ways. Research into the cognitive correlates of pupil-size is currently enjoying something of a renaissance, but even now there have been very few studies that have explored the role of mnemonic processes. The topic is still in its infancy and several important issues remain unresolved, with some key methodological issues yet to be refined.

The eye-tracker outputs other data, so other aspects which might be interesting to analyse include latency to peak pupil-size, in order to understand more about the timecourse of the underlying neurophysiological processes and further characterise the PONE. The data collected during the course of this thesis could be analysed as waveforms in a manner similar to the analysis of ERPs. This would allow inspection of grand-average waveforms and selection of smaller time windows within the 2000ms trials for analysis if there appears to be a consistent pattern of response.

Then a mean ‘amplitude’ measure from a specific time window could be extracted. For example, Kuipers and Thierry (2011) selected an epoch within each trial from -100 to 850ms, where 0ms represents stimulus presentation, baseline-corrected to the 100ms prestimulus section, and applied a 10Hz low-pass filter offline prior to measurement of mean pupil amplitude. Other researchers report filtering pupil-size data using a 10Hz low-pass filter (e.g., Hupé et al., 2009) as cognition-induced changes in pupil size are of a lower frequency, but some sources of noise (e.g., from estimation of pupil-size) produce artifactual changes at a higher frequency. The application of artifact-detection might also give cleaner data with which to work (e.g., Hupé et al., 2009).

One way in which the findings of this thesis could be taken forward would be to explore the effects of other types of memory strength manipulations on the PONE. If, as argued above, the PONE essentially reflects a strength of memory signal, then other manipulations of memory strength should also result in a “graded” pupil-size. For example the use of established mnemonic strategies, such as visualisation/imagery for half the stimuli at learning, should enhance memory, facilitating a much stronger recollection at recognition, and therefore a bigger pupil. Additionally it would be interesting to investigate the effects on the PONE of manipulations which have been shown to dissociate behavioural measures of familiarity and recollection in the ERP literature. For example, Yovel and Paller’s (2004) unknown faces with spoken occupations design could be adapted for use with words by presenting related information with each item at learning and asking participants to recall the additional information at recognition. This would allow separation of items remembered with different degrees of clarity for analysis.

Given the promising preliminary finding that a similar pattern emerges in the PONE as in ERP correlates of recognition memory, further concurrent recordings could be made to investigate whether the PONE responses to ‘strength of memory’ manipulations were matched by variations in the magnitude of the ERP old/new effects at 300-500ms and 500-700ms. Once the time-course of corresponding effects had been clarified using ERPs, a technique with better spatial resolution could be applied, such as EEG source localisation or fMRI, making links with what is already known about the neural substrate of recognition memory and of the pupillary control system.

In recent years, researchers have shown that it is possible to record from single neurons in conscious human brains (see for example Quiñ Quiroga’s work on the Jennifer Anniston neuron, 2008; 2010; Rutishauser et al., 2008). If the opportunity arose to test recognition memory in patients undergoing awake brain surgery whilst also measuring pupil-size and directly recording the activity of Locus Coeruleus (LC) neurons, this would provide valuable information confirming the hypothesised link between LC activity and pupil-size in humans during cognitive tasks. It would also provide information on whether the LC increases phasic firing when participants are correctly recognising old items compared to new items, or whether this increased dilation arises from a different part of the brain.

It is noted that no studies so far deal directly with pupil-size during visual word recognition in amnesic patients, reflecting a gap in the literature. In terms of pupillometry research, this suggests that future work needs to pay more attention to what happens to the apparently conscious recognition-related PONE in a patient who can’t make an explicit recognition decision. These ideas were introduced in Chapter 5, and it was suggested that the PONE might be suitable as a means for

detecting malingered memory-impairment. However, the only way to truly determine the PONE's utility would be by establishing its parameters within a genuine patient population in a standard recognition memory task.

As well as exploring the PONE in amnesic patients, memory could be manipulated in healthy participants using Transcranial Magnetic Stimulation (TMS), which has been shown to impair some types of memory (e.g., Prime, Vesia, & Crawford, 2008), and enhance others (Kirschen, Davis-Ratner, Jerde, Schraedley-Desmond, & Desmond, 2006). Machizawa, Kalla, Walsh and Otten (2010) found that TMS applied to the left or right inferior frontal gyrus of the prefrontal cortex during the learning phase affected performance on a recognition memory test fifteen minutes later. Turriziani et al. (2008) found that familiarity was impaired in a test of recognition after TMS was used to stimulate the right and left dorsolateral prefrontal cortex (DLPFC) prior to encoding, and that recollection was impaired after stimulation of the right DLPFC. Clearly it is possible to influence behavioural performance measures on a recognition test and it would be fascinating to see whether this temporary impairment also extends to the PONE.

8. Bibliography

- Aboyoun, D.C., & Dabbs, J.N. (1998). The Hess pupil dilation findings: Sex or novelty? *Social Behavior and Personality*, 26(4), 415-419.
doi:10.2224/sbp.1998.26.4.415
- Achilles, E.M. (1920). *Experimental studies in recall and recognition*. New York: G.E. Stechert & Co. Retrieved 14 October 2011 from
<http://www.archive.org/details/experimentalstud00achi>
- Aggleton, J.P., & Brown, M.W. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behavioral and Brain Sciences*, 22(3), 425–489.
- Ahern, S.K., & Beatty, J. (1979). Pupillary responses during information processing vary with scholastic aptitude test scores. *Science*, 205(4412), 1289-1292.
doi:10.1126/science.472746
- Ahern, S.K., & Beatty, J. (1981). Physiological evidence that demand for processing capacity varies with intelligence. In M.P. Friedman, J.P. Das, & N. O'Connor (Eds.), *Intelligence and learning* (pp. 121-128). New York: Plenum.
- Allan, K., & Rugg, M.D. (1997). An event-related potential study of explicit memory on tests of cued recall and recognition. *Neuropsychologia*, 35(4), 387-397.
doi:10.1016/S0028-3932(96)00094-2
- Allan, K., Wilding, E.L., & Rugg, M.D. (1998). Electrophysiological evidence for dissociable processes contributing to recollection. *Acta Psychologica*, 98(2-3), 231–252. doi:10.1016/S0001-6918(97)00044-9
- Allen, J., & Mertens, R. (2009). Limitations to the detection of deception: True and false recollections are poorly distinguished using an event-related potential procedure. *Social Neuroscience*, 4(6), 473-490.
doi:10.1080/17470910802109939
- Allen, R., & Reber, A.S. (1980). Very long-term memory for tacit knowledge. *Cognition*, 8(2), 175-185. doi:10.1016/0010-0277(80)90011-6
- Ally, B.A., Waring, J.D., Beth, E.H., McKeever, J.D., Milberg, W.P., & Budson, A.E. (2008). Aging memory for pictures: Using high-density event-related potentials to understand the effect of aging on the picture superiority effect. *Neuropsychologia*, 46(2), 679-689.
doi:10.1016/j.neuropsychologia.2007.09.011

- American Clinical Neurophysiology Society (2006). Guideline Five: Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, 23(2), 107-110.
- American Electroencephalographic Society (1991). Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, 8, 200-202.
- American Electroencephalographic Society (1994). Guideline thirteen: Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, 11, 111-113.
- Andreassi, J.L. (2000). *Psychophysiology: Human behavior and physiological response*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Ardolf, B.R., Denney, R.L., & Houston, C.M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist*, 21(6), 899-916. doi:10.1080/13825580600966391
- Aston-Jones, G., Chiang, C., & Alexinsky, T. (1991). Discharge of noradrenergic locus coeruleus neurons in behaving rats and monkeys suggests a role in vigilance. *Progress in Brain Research*, 88, 501-20.
- Aston-Jones, G., & Cohen, J.D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, 28, 403-450. doi:10.1146/annurev.neuro.28.061604.135709
- Aston-Jones, G., Foote, S.L., & Bloom, F.E. (1984). Anatomy and physiology of locus coeruleus neurons: Functional implications. In M. Ziegler, & C.R. Lake (Eds.), *Norepinephrine: Frontiers of clinical neuroscience* (pp. 92-116). Baltimore: Williams and Wilkins.
- Aston-Jones, G., Rajkowski, J., & Kubiak, P. (1997). Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. *Neuroscience*, 80(3), 697-715. doi:10.1016/S0306-4522(97)00060-2
- Aston-Jones, G., Rajkowski, J., Kubiak, P., & Alexinsky, T. (1994). Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *The Journal of Neuroscience*, 14(7), 4467-4460.

- Aston-Jones, G., Rajkowski, J., Lu, W., Zhu, Y., Cohen, J.D., & Morecraft, R.J. (2002). Prominent projections from the orbital prefrontal cortex to the locus coeruleus in monkey. *Society for Neuroscience Abstracts*, 28, 86-89.
- Atkinson, R.C., & Juola, J.F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.) *Attention and performance IV* (pp. 583-612). New York: Academic Press.
- Atkinson, R.C., & Juola, J.F. (1974). Search and decision processes in recognition memory. In D.H. Krantz, R.C. Atkinson, R.D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology. I: Learning, memory and thinking* (pp. 243-293). San Francisco, CA: Freeman.
- Axelrod, B.N., Barlow, A., & Paradee, C. (2010). Evaluation of the WMS-III rarely missed index in a naïve clinical sample. *Clinical Neuropsychologist*, 24(1), 95-102. doi:10.1080/13854040902960083
- Backs, R.W., & Walrath, L.C. (1992). Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. *Applied Ergonomics*, 23(4), 243-254. doi:10.1016/0003-6870(92)90152-L
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423. doi:10.1016/S1364-6613(00)01538-2
- Baddeley, A.D., & Hitch, G.J.L. (1974). Working memory. In G.A. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89), New York: Academic Press.
- Bailey, B.P., Busbey, C.W., & Iqbal, S.T. (2007). TAPRAV: An interactive analysis tool for exploring workload aligned to models of task execution. *Interacting with Computers*, 19(3), 314-329. doi:10.1016/j.intcom.2007.01.004
- Bailey, B.P., & Iqbal, S.T. (2008). Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management. *ACM Transactions on Computer-Human Interaction*, 14(4), 1-28. doi:10.1145/1314683.1314689
- Bailey, D.L., Townsend, D.W., Valk, P.E., & Maissey, M.N. (2005). *Positron emission tomography: Basic sciences*. London: Springer.
- Balota, D.A., & Neely, J.H. (1980). Text expectancy and word-frequency effects in recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory*, 6(5), 576–587. doi:10.1037/0278-7393.6.5.576

- Banks, W.P. (1970). Signal detection theory and human memory. *Psychological Bulletin*, 74(2), 81-99. doi:10.1037/h0029531
- Barbur, J.L., Wolf, J., & Lennie, P. (1998). Visual processing levels revealed by response latencies to changes in different visual attributes. *Proceedings of the Royal Society of London, Series B*, 265(1412), 2321–2325. doi:10.1098/rspb.1998.0578
- Barlow, J.D. (1969). Pupillary size as a index of preference in political candidates. *Perceptual & Motor Skills*, 28, 927-932. doi:10.2466/pms.1969.28.2.587
- Barlow, J.D. (1970). Pupillary size as an index of preference. *Perceptual & Motor Skills*, 31(1), 331-336. doi:10.2466/pms.1970.31.1.331
- Barreto, A., Zhai, J., Rishe, N., & Gao, Y. (2007). Significance of pupil diameter measurements for the assessment of affective state in computer users. In E. Elleithy (Ed.), *Advances and innovations in systems, computing sciences and software engineering* (pp. 59-64). Dordrecht, The Netherlands: Springer. doi:10.1007/978-1-4020-6264-3_11
- Bartholomew, A. (1483). *De proprietatibus rerum, Liber V*. Nuremberg, Germany: Anton Koberger. (Original work published around 1240). Retrieved 25 August 2011 from <http://catalog.hathitrust.org/Record/001906553>
- Bauer, L.M., Olheiser, E.L., Altarriba, J., & Landi, N. (2009). Word type effects in false recall: Concrete, abstract, and emotion word critical lures. *American Journal of Psychology*, 122(4), 469-481.
- Baxendale, S. (2004). Memories aren't made of this: Amnesia at the movies. *British Medical Journal*, 329(7480), 18-25. doi:10.1136/bmj.329.7480.1480
- Beatty, J. (1982a). Phasic not tonic pupillary responses vary with auditory vigilance performance. *Psychophysiology*, 19(2), 167-172. doi:10.1111/j.1469-8986.1982.tb02540.x
- Beatty, J. (1982b). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91, 276–292. doi:10.1037/0033-2909.91.2.276
- Beatty, J. (1986). The pupillary system. In G.H. Coles, E. Donchin, & S.W. Porges (Eds.), *Psychophysiology: Systems, process, and applications* (pp. 43-50). New York: Guilford Press.
- Beatty, J. (1988). Pupillometric signs of selective attention in man. In G.C. Galbraith, M.L. Kietzman, & E. Donchin (Eds.), *Neurophysiology and psychophysiology: Experimental and clinical applications* (pp. 138-143). Hillsdale, NJ: Erlbaum.

- Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science*, 5, 371-372.
- Beatty, J., & Lucero-Wagoner, B. (1975). Pupillary measurement of sensory and decision processes in a signal-detection task. Paper presented at *The Meeting of the Psychonomic Society*. Denver: Psychonomic Society.
- Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J.T. Cacioppo, L.G. Tassinary, & G.G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142-162). USA: Cambridge University Press.
- Beatty, J., & Wagoner, B.L. (1976). Pupillometric assessment of signal analysis and decision processes in acoustic signal detection. Unpublished manuscript, University of California at Los Angeles, CA, USA.
- Beatty, J., & Wagoner, B.L. (1978). Pupillometric signs of brain activation vary with level of cognitive processing. *Science*, 199(4334), 1216-1218. doi:10.1126/science.628837
- Becker, W., Iwase, K., Jürgens, R., & Kornhuber, H.H. (1976). Brain potentials preceding slow and rapid hand movements. In W.C. McCallum, & J.R. Knott (Eds.), *The responsive brain* (pp. 99-102). Bristol: Wright.
- Bender, W.R.G. (1933). The effect of pain and emotional stimuli and alcohol upon pupillary reflex activity. *Psychological Monographs*, 44(2), 1-32.
- Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology*, 88(1), 131-151. doi:10.1037/0021-9010.88.1.131
- Bentin, S., & McCarthy, G. (1994). The effects of immediate stimulus repetition on reaction time and event-related potentials in tasks of different complexity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20(1), 130-149.
- Bentin, S., & Peled, B. (1990). The contribution of task-related factors to ERP repetition effects at short and long lags. *Memory & Cognition*, 18(4), 359-366. doi:10.3758/BF03197125
- Bernard, C. (1852). Sur les effets de la section de la portion céphalique du grand sympathique. *Chronicles of the Royal Society of Biologists, Paris*, 4, 168-170.
- Bernbach, H.A. (1967). Decision processes in memory. *Psychological Review*, 74(6), 462-480. doi:10.1037/h0025132

- Bernhardt, P.C., Dabbs, J.M., & Riad, J.K. (1996). Pupillometry system for use in social psychology. *Behavior Research Methods*, 28(1), 61–66.
doi:10.3758/BF03203637
- Bernick, N., Altman, F., & Mintz, D.L. (1972). Pupil responses of addicts in treatment to drug culture argot: II. Responses during verbalization of visually presented words. *Psychonomic Science*, 28(2), 81-82.
- Bernick, N., Kling, A., & Borowitz, G. (1971). Physiologic differentiation of sexual arousal and anxiety. *Psychosomatic Medicine*, 33(4), 341-352.
- Berridge, C.W., & Waterhouse, B.D. (2003). The locus coeruleus- noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research Review*, 42, 33–84. doi:10.1016/S0165-0173(03)00143-7
- Berrien, F.K., & Huntington, G.H. (1943). An exploratory study of pupillary responses during deception. *Journal of Experimental Psychology*, 32(5), 443-449.
doi:10.1037/h0063488
- Berry, D., & Dienes, Z.P. (1993). *Implicit learning: Theoretical and empirical issues*. Hove: Lawrence Erlbaum Associates Ltd.
- Birren, J.E., Casperson, R.C., & Botwinick, J. (1950). Age changes in pupil-size. *Journal of Gerontology*, 5(3), 216-221. doi:10.1093/geronj/5.3.216
- Bitsios, P., Philpott, A., Langley, R.W., Bradshaw, C.M., & Szabadi, E. (1999). Comparison of the effects of diazepam on the fear-potentiated startle reflex and the fear-inhibited light reflex in man. *Journal of Psychopharmacology*, 13(3), 226-234. doi:10.1177/026988119901300303
- Bitsios, P., Prettyman, R., & Szabadi, E. (1996a). Changes in autonomic function with age: A study of pupillary kinetics in healthy young and old people. *Age & Ageing*, 25(6), 432-438. doi:10.1093/ageing/25.6.432
- Bitsios, P., Szabadi, E., & Bradshaw, C.M. (1996b). The inhibition of the pupillary light reflex by the threat of an electric shock: A potential laboratory model of human anxiety. *Journal of Psychopharmacology*, 10(4), 279-287.
doi:10.1177/026988119601000404
- Bitsios, P., Szabadi, E., & Bradshaw, C.M. (2002). Relationship of the 'fear-inhibited light reflex' to the level of state/trait anxiety in healthy subjects. *International Journal of Psychophysiology*, 43(2), 177-184. doi:10.1016/S0167-8760(01)00173-8

- Botvinick, M., Braver, T., Barch, D., Carter, C., & Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108(3), 624-652. doi:10.1037//0033-295X.108.3.624
- Bouma, H. (1962). Size of the static pupil as a function of wavelength and luminosity of the light incident on the human eye. *Nature*, 193, 690–691. doi:10.1038/193690a0
- Bouma, H., & Baghuis, L.C.J. (1971). Hippus of the pupil: Periods of slow oscillations of unknown origin. *Vision Research*, 11(11), 1345-1351. doi:10.1016/0042-6989(71)90016-2
- Bourget, D., & Whitehurst, L. (2007). Amnesia and crime. *Journal of the American Academy of Psychiatry and the Law*, 35(4), 469–480.
- Bowles, B., Crupi, C., Mirsatarri, S.M., Pigott, S.E., Parrent, A.G., Pruessner, J.C., ... Kohler, S. (2007). Impaired familiarity with preserved recollection after anterior temporal lobe resection that spares the hippocampus. *Proceedings for the National Academy of Sciences*, 104(41), 16382-16387. doi:10.1073/pnas.0705273104
- Bowles, B., Crupi, C., Pigott, S., Parrent, A., Wiebe, S., Janzen, L., & Kohler, S. (2010). Double dissociation of selective recollection and familiarity impairments following two surgical treatments for temporal lobe epilepsy. *Neuropsychologia*, 48(9): 2640–2647. doi:10.1016/j.neuropsychologia.2010.05.010
- Bradley, J.C., Bentley, K.C., Mughal, A.I., Bodhireddy, H., Young, R.S., & Brown, S.M. (2010). The effect of gender and iris color on the dark-adapted pupil diameter. *Journal of Ocular Pharmacology and Therapeutics*, 26(4), 335-40. doi:10.1089/jop.2010.0061
- Bradley, M.T., & Janisse, M.P. (1979). Pupil-size and lie detection: The effect of certainty on detection. *Psychology*, 16(4), 33-39.
- Bradley, M.T., & Janisse, M.P. (1981). Accuracy demonstrations, threat, and the detection of deception: Cardiovascular, electrodermal, and pupillary measures. *Psychophysiology*, 18(3), 307-315. doi:10.1111/j.1469-8986.1981.tb03040.x
- Bradshaw, J. (1967). Pupil-size as a measure of arousal during information processing. *Nature*, 216(5114), 515-516. doi:10.1038/216515a0
- Bradshaw, J.L. (1968a). Pupil-size and problem solving. *Quarterly Journal of Experimental Psychology*, 20(2), 116-122. doi:10.1080/14640746808400139

- Bradshaw, J.L. (1968b). Pupillary changes and reaction time with varied stimulus uncertainty. *Psychonomic Science*, 13(2), 69-70.
- Bradshaw, J.L. (1969). Background light intensity and the pupillary response in a reaction time task. *Psychonomic Science*, 14(6), 271-272.
- Bradshaw, J.L. (1970). Pupil-size and drug state in a reaction time task. *Psychonomic Science*, 18(2), 112-113.
- Brainerd, C.J., & Reyna, V.F. (2002). Recollection rejection: How children edit their false memories. *Developmental Psychology*, 38(1), 156-172.
doi:10.1037/0012-1649.38.1.156
- Brainerd, C.J., Wright, R., Reyna, V.F., & Mojardin, A.H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 27(2), 307-327. doi:10.1037/0278-7393.27.2.307
- Brontë, C. (1946). *Jane Eyre*. M. Lamberton Becker, (Ed.). Cleveland, NY: The World Publishing Co. (Original work published 1847).
- Brown, A.A., & Bodner, G.E. (2011). Re-examining dissociations between remembering and knowing: Binary judgments vs. independent ratings. *Journal of Memory and Language*, 65(2), 98-108.
doi:10.1016/j.jml.2011.04.003
- Brown, G.G., Kinderman, S.S., Siegle, G.J., Granholm, E., Wong, E.C., & Buxton, R.B. (1999). Brain activation and pupil response during covert performance of the Stroop colour word task. *Journal of the International Neuropsychological Society*, 5(4), 308-319.
- Browndyke, J.N., Paskavitz, J., Sweet, L.H., Cohen, R.A., Tucker, K.A., Welsh-Bohmer, K.A., ... Schmechel, D.E. (2008). Neuroanatomical correlates of malingered memory-impairment: Event-related fMRI of deception on a recognition memory task. *Brain Injury*, 22(6), 481-489.
doi:10.1080/02699050802084894
- Brozinsky, C.J., Yonelinas, A.P., Kroll, N.E.A., & Ranganath, C. (2005). Lag-sensitive repetition suppression effects in the anterior parahippocampal gyrus. *Hippocampus*, 15(5), 557-561. doi:10.1002/hipo.20087
- Brunia, C.H.M., & Vingerhoets, A.J.J.M. (1980). CNV and EMG preceding a plantar flexion of the foot. *Biological Psychology*, 11(3-4), 181-191.
- Buckner, R.L., Goodman, J., Burock, M., Rotte, M., Koutstaal, W., Schacter, D., ... Dale, A.M. (1998). Functional-anatomic correlates of object priming in

- humans revealed by rapid presentation event-related fMRI. *Neuron*, 20(2), 285–296. doi:10.1016/S0896-6273(00)80456-0
- Budge, J., & Waller, A. (1851). Action de la portion cervicale du nerf grand sympathique et d'une portion de la moelle epinière sur la dilatation de la pupille. *Gazette Médicale de Paris: Journal de Médecine et des Sciences Accessoires*, 3(6), 648-649.
- Bull, R., & Shead, G. (1979). Pupil-dilation, sex of stimulus, and age and sex of observer. *Perceptual & Motor Skills*, 49(1), 27-30. doi:10.2466/pms.1979.49.1.27
- Cacioppo, J.T., & Tassinary, L.G. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, 45(1), 16-28.
- Cacioppo, J.T., Tassinary, L.G., & Berntson, G.G. (2000). Psychophysiological science. In J.T. Cacioppo, L.G. Tassinary, & G.G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 3-23). New York: Cambridge University Press.
- Cain, B. (2007). *A review of the mental workload literature* (NATO Report No. RTO-TR-HFM-121-Part-II). Toronto, Canada: Defence Research and Development.
- Carver, R.P. (1971). Pupil dilation and its relationship to information processing during reading and listening. *Journal of Applied Psychology*, 55(2), 126-134. doi:10.1037/h0030664
- Casey, B.J., & de Haan, M. (2002). Introduction: New methods in developmental science. *Developmental Science*, 5(3), 265-267. doi:10.1111 /1467-7687.00365
- Chan, C. (1992). *Implicit cognitive processes: Theoretical issues and applications in computer systems design*. (Unpublished doctoral dissertation). University of Oxford, Oxford.
- Chapman, C.R., Oka, S., Bradshaw, D.H., Jacobson, R.C., & Donaldson, G.W. (1999). Phasic pupil dilation response to noxious stimulation in normal volunteers: Relationship to brain evoked potentials and pain report. *Psychophysiology*, 36(1), 44-52. doi:10.1017/S0048577299970373
- Chatham, C.H, Frank, M.J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences*, 106(14), 5529-5533, 2009. doi:10.1073/pnas.0810002106

- Chelazzi, L., Miller, E.K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, 363, 345–347. doi:10.1038/363345a0
- Cherry, K.E., & Stadler, M.A. (1995). Implicit learning of a non-verbal sequence in younger and older adults. *Psychology & Aging*, 10(3), 379-394. doi:10.1037/0882-7974.10.3.379
- Cima, M., Nijman, H., Merckelbach, H., Kremer, K., & Hollnack, S. (2004). Claims of crime-related amnesia in forensic patients. *International Journal of Law & Psychiatry*, 27(3), 215–221. doi:10.1016/j.ijlp.2004.03.007
- Clark, J.W. (1885). Civilisation and eyesight. *Nature*, 31(802), 433–434. doi:10.1038/031433d0
- Clark, K.B., Krah, S.E., Smith, D.S., & Jensen, R.A. (1995). Post-training unilateral vagal stimulation enhances retention performance in the rat. *Neurobiology of Learning & Memory*, 63(3), 213-216. doi:10.1006/nlme.1995.1024
- Clark, K.B., Naritoku, D.K., Smith, D.C., Browning, R.A., & Jensen, R.A. (1999). Enhanced recognition memory following vagus nerve stimulation in human subjects. *Nature Neuroscience*, 2, 94-98. doi:10.1038/4600
- Clark, S.E., & Gronlund, S.D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, 3(1), 37–60. doi:10.3758/BF03210740
- Clark, W.R. (1975). A comparison of pupillary response, heart rate, and GSR during deception. Paper presented at *The Annual Meeting of the Midwestern Psychological Association*. Chicago: Midwestern Psychological Association.
- Clark, W.R., & Ertas, M.A. (1975). A comparison of pupillary reactions to visual and auditory stimuli in a test of preferences for presidential candidates. *Journal Supplement Abstract Service Catalog of Selected Documents in Psychology*, 12, 20-25.
- Clark, W.R., & Johnson, D.A. (1970). Effects of instructional set on pupillary responses during a short-term memory task. *Journal of Experimental Psychology*, 85(2), 315-317. doi:10.1037/h0029543
- Clayton, E.C., Rajkowski, J., Cohen, J.D., & Aston-Jones, G. (2004). Phasic activation of monkey locus ceruleus neurons in a forced-choice task. *Journal of Neuroscience*, 24(44), 9914–9920. doi:10.1523/JNEUROSCI.2446-04.2004

- Cleeremans, A., & Dienes, Z. (2008). Computational models of implicit learning. In R. Sun (Ed.), *The Cambridge handbook of computational modeling* (pp. 396-421). Cambridge, UK: Cambridge University Press.
- Cohen, J.D., Aston-Jones, G., & Gilzenrat, M.S. (2004). A systems-level perspective on attention and cognitive control: Guided activation, adaptive gating, conflict monitoring, and exploitation versus exploration. In M.I. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 71–90). New York: Guilford Press.
- Cohen, J.D., Botvinick, M., & Carter, C.S. (2000). Anterior cingulate and prefrontal cortex: Who's in control? *Nature Neuroscience*, 3, 421-423.
doi:10.1038/74783
- Cohn, M., Moscovitch, M., & Davidson, P.S.R. (2010). Double dissociation between familiarity and recollection in Parkinson's disease as a function of encoding tasks. *Neuropsychologia*, 48(14), 4142-4147.
doi:10.1016/j.neuropsychologia.2010.10.013
- Coles, M.G.H., & Strayer, D.L. (1985). The psychophysiology of the cardiac cycle time effect. In J.F. Orlebeke, G. Mulder, & L.J.P. van Doornen (Eds.), *Psychophysiology of cardiovascular control: Models, methods and data* (pp. 517-534). New York: Plenum Press.
- Colman, F.D., & Paivio, A. (1969). Pupillary response and galvanic skin response during an imagery task. *Psychonomic Science*, 16(6), 296-297.
- Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 33A, 497-505. Accessed at http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm
- Conati, C., & Merten, C. (2007). Eye-tracking for user modeling in exploratory learning environments: An empirical evaluation. *Knowledge-Based Systems*, 20(6), 557-574. doi:10.1016/j.knosys.2007.04.010
- Conway, C.A., Jones, B.C., DeBruine, L.M., Little, A.C., & Sahraie, A. (2008). Transient pupil constrictions to faces are sensitive to orientation and species. *Journal of Vision*, 8(3), 17, 1-11. doi:10.1167/8.3.17
- Cottingham, M.E., & Boone, K.B. (2010). Non-credible language deficits following mild traumatic brain injury. *The Clinical Neuropsychologist*, 24(6), 1006-1025.
doi:10.1080/13854046.2010.481636
- Craik, F.I.M., Govoni, R., Naveh-Benjamin, M., & Anderson, N.D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, 125(2), 159-180.

- Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal learning & Verbal Behaviour*, 11(6), 671-684. doi:10.1016/S0022-5371(72)80001-X
- Craik, F.I.M., & McDowd, J.M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 13(3), 474-479.
- Criss, A.H. (2006). The consequences of differentiation in episodic memory: Similarity and the strength based mirror effect. *Journal of Memory & Language: Special Issue on Computational Models of Memory*, 55(4), 461-478. doi:10.1016/j.jml.2006.08.003
- Curran, T. (1999). The electrophysiology of incidental and intentional retrieval: ERP old/new effects in lexical decision and recognition memory. *Neuropsychologia*, 37(7), 771-785. doi:10.1016/S0028-3932(98)00133-X
- Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, 28(6), 923-938.
- Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia*, 42(8), 1088-1106. doi:10.1016/j.neuropsychologia.2003.12.011
- Curran, T., & Cleary, A.M. (2003). Using ERPs to dissociate recollection from familiarity in picture recognition. *Cognitive Brain Research*, 15(2), 191-205. doi:10.1016/S0926-6410(02)00192-1
- Curran, T., DeBuse, C., Woroch, B., & Hirshman, E. (2006). Combined pharmacological and electrophysiological dissociation of familiarity and recollection. *The Journal of Neuroscience*, 26(7), 1979-1985. doi:10.1523/JNEUROSCI.5370-05.2006
- Curran, T., & Doyle, J. (2011). Picture superiority doubly dissociates the ERP correlates of recollection and familiarity. *Journal of Cognitive Neuroscience*, 23(5), 1247-1262. doi:10.1162/jocn.2010.21464
- Curran, T., Schacter, D.L., Johnson, M.K., & Spinks, R. (2001). Brain potentials reflect behavioral differences in true and false recognition. *Journal of Cognitive Neuroscience*, 13, 201-216. doi:10.1162/089892901564261
- Curran, T., Tepe, K.L., & Piatt, C. (2006). ERP explorations of dual processes in recognition memory. In H.D. Zimmer, A. Mecklinger, & U. Lindenberger (Eds.), *Binding in human memory: A neurocognitive approach* (pp. 467-492). Oxford: Oxford University Press.

- Curran, T., Tucker, D.M., Kutas, M., & Posner, M.I. (1993). Topography of the N400: Brain electrical activity reflecting semantic expectation. *Electroencephalography & Clinical Neurophysiology*, 88(3), 188–209.
- Dabbs, J.M., & Milun, R. (1999). Pupil dilation when viewing strangers: Can testosterone moderate prejudice? *Social Behavior & Personality*, 27(3), 297–301.
- Dahlberg, J. (2010). *Eye tracking with eye glasses*. (Unpublished masters dissertation). Umea University, Sweden.
- Dain, S.J., Cassimaty, V.T., & Psarakis, D.T. (2004). Differences in FM100-Hue test performance related to iris colour may be due to pupil-size as well as presumed amounts of macular pigmentation. *Clinical & Experimental Optometry*, 87(4-5), 322-325. doi:10.1111/j.1444-0938.2004.tb05061.x
- Daly, R.M. (1966). *Pupillary size and its relationship to the problem solving process*. (Unpublished doctoral dissertation). Loyola University, Chicago.
- Darwin, C.R. (1872). *The expression of the emotions in man and animals*. London: John Murray.
- Daugman, J. (2003). The importance of being random: statistical principles of iris recognition. *Pattern Recognition*, 36(2), 279–291. doi:10.1016/S0031-3203(02)00030-4
- Davies, L.N., Wolffsohn, J.S., & Gilmartin, B. (2005). Cognition, ocular accommodation, and cardiovascular function in emmetropes and late-onset myopes. *Investigative Ophthalmology & Visual Science*, 46(5), 1791-1796. doi:10.1167/iov.04-0986
- de Balzac, H. (2000). *Le curé du village [The village rector]*. (K. Prescott Wormeley, Trans.). Cirencester: The Echo Library. (Original work published 1841).
- de Greef, T., Lafeber, H., van Oostendorp, H., & Lindenberg, J. (2009). Eye movement as indicators of mental workload to trigger adaptive automation. *Foundations of augmented cognition: Neuroergonomics and operational neuroscience* (pp. 219-228): Springer-Verlag.
- Deese, J. (1960). Frequency of usage and number of words in free recall: The role of association. *Psychological Review*, 7, 337-344. doi: 10.2466/PR.7.6.337-344
- Dehais, F., Causse, M., & Pastor, J. (2008). Embedded eye tracker in a real aircraft: new perspectives on pilot/aircraft interaction monitoring. In Proceedings from

- The 3rd International Conference on Research in Air Transportation*. Fairfax, USA: Federal Aviation Administration.
- DeLaunay, J. (1949). A note on the photo-pupil reflex. *Journal of the Optical Society of America*, 39(5), 364-367. doi:10.1364/JOSA.39.000364
- Denney, R.L. (2007). Assessment of malingering in criminal forensic neuropsychological settings. In K.B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 428–452). New York: Guilford.
- Dennis, S., & Humphreys, M.S. (2001). A context noise model of episodic word recognition. *Psychological Review*, 108(2), 452–478. doi:10.1037/0033-295X.108.2.452
- DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74-118. doi:10.1037/0033-2909.129.1.74
- Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences of the United States of America*, 93(24), 13494–13499.
- Diana, R., Reder, L.M., Arndt, J., & Park, H. (2006). Models of recognition: A review of arguments in favor of a dual process account. *Psychonomic Bulletin & Review*, 13(1), 1-21.
- Dien, J. (1998). Issues in the application of the average reference: Review, critiques, and recommendations. *Behavior Research Methods*, 30(1), 34–43. doi:10.3758/BF03209414
- Dienes, Z., & Altmann, G. (1997). Transfer of implicit knowledge across domains: How implicit and how abstract? In D. Berry (Ed.), *How implicit is implicit learning?* (pp 107-123). Oxford: Oxford University Press.
- Dienes, Z., Altmann, G., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 21(5), 1322-1338. doi:10.1037/0278-7393.21.5.1322
- Dienes, Z., & Berry, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin & Review*, 4(1), 3-23.
- Dienes, Z., Broadbent, D.E., & Berry, D.C. (1991). Implicit and explicit knowledge bases in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 17(5), 875-887.

- Dienes, Z., & Fahey, R. (1998). The role of implicit memory in controlling a dynamic system. *Quarterly Journal of Experimental Psychology A*, 51(3), 593-614.
- Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research*, 69(5-6), 338-351. doi:10.1007/s00426-004-0208-3
- Diller, D.E., Nobel, P.A., & Shiffrin, R.M. (2001). An ARC–REM model for accuracy and response time in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 27(2), 414-435. doi:10.1037/0278-7393.27.2.414
- Dionisio, D.P., Granholm, E., Hillix, W.A., & Perrine, W.F. (2001). Differentiation of deception using pupillary responses as an index of cognitive processing. *Psychophysiology*, 38(2), 205-211.
- Donaldson, D.I., Allan, K., & Wilding, E.L. (2002). Fractionating episodic memory retrieval using event-related potentials. In A. Parker, E.L. Wilding, & T. Bussey (Eds.), *The cognitive neuroscience of memory: Encoding and retrieval* (pp.39-58). Hove: Psychology Press.
- Donaldson, D.I., Petersen, S.E., Ollinger, J.M., & Buckner, R.L. (2001). Dissociating state and item components of recognition memory using fMRI. *NeuroImage*, 13, 129-142. doi:10.1006/nimg.2000.0664
- Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition*, 24(4), 523-533.
- Dooley, R.P., & Lehr, D.J. (1967). Critique of a pupillary response experiment. *Perceptual & Motor Skills*, 25(2), 603-604. doi:10.2466/pms.1967.25.2.603
- Dosher, B.A. (1984). Degree of learning and retrieval speed: Study time and multiple exposures. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 10(4), 541–574. doi:10.1037/0278-7393.10.4.541
- Doughty, M.J. (2001). *Ocular pharmacology and therapeutics: A primary care guide*. Oxford: Butterworth-Heinemann.
- Duarte, A., Ranganath, C., Winward, L., Hayward, D., & Knight, R.T. (2004). Dissociable neural correlates for familiarity and recollection during the encoding and retrieval of pictures. *Cognitive Brain Research*, 18(3), 255-272. doi:10.1016/j.cogbrainres.2003.10.010
- Duchowski, A.T. (2007). *Eye tracking methodology: Theory and practice* (2nd ed.). London: Springer.

- Dunn, J.A. (2004). Remember-Know: A matter of confidence. *Psychological Review*, 111(2), 524-542.
- Düzel, E., Vargha-Khadem, F., Heinze, H.J., & Mishkin, M. (2001). Brain activity evidence for recognition without recollection after early hippocampal damage. *Proceedings of the National Academy of Sciences of the United States of America*, 98(14), 8101-8106. doi:10.1073/pnas.131205798
- Düzel, E., Yonelinas, A.P., Mangun, G.R., Heinze, H.J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Sciences of the United States of America*, 94(11), 5973-5978.
- Einhauser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Science of the United States of America*, 105(5), 1704-1709. doi:10.1073/pnas.0707727105
- Ekman, I., Poikola, A., Mäkräinen, M., Takala, T., & Hämäläinen, P. (2008a). Voluntary pupil-size change as control in eyes only interaction. In *Proceedings from The 2008 Symposium on Eye Tracking Research & Applications* (pp. 115-188). New York, NY: ACM Press. doi:10.1145/1344471.1344501
- Ekman, I.M., Poikola, A.W., & Mäkräinen, M.K. (2008b). Invisible eni: Using gaze and pupil-size to control a game. In *Proceedings from The CHI '08 Extended Abstracts on Human Factors in Computing Systems* (pp. 3135-3140). New York, NY: ACM Press. doi:10.1145/1358628.1358820
- Elaad, E. (2009). Lie-detection biases among male police interrogators, prisoners, and laypersons. *Psychological Reports*, 105(3), 1047-1056. doi:10.2466/pr0.105.F.1047-1056
- Eldridge, L.L., Knowlton, B.J., Furmanski, C.S., Bookheimer, S.Y., & Engel, S.A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience*, 3(11), 1149-1152. doi:10.1038/80671
- Elshtain, E.L., & Schaefer, T.Jr. (1968). Effects of storage load and word frequency on pupillary responses during short-term memory. *Psychonomic Science*, 12(4), 143-144.
- Engineer, N.D., Riley, J.R., Seale, J.D., Vrana, W.A., Shetake, J.A., Sudanagunta, S.P., . . . Kilgard, M.P. (2011). Reversing pathological neural activity using targeted plasticity. *Nature*, 470, 101-106. doi:10.1038/nature09656

- Engelhard, I.M., Merckelbach, H., & van den Hout, M.A. (2003). The Guilty Knowledge Test and the modified Stroop task in detection of deception: An exploratory study. *Psychological Reports*, 92(2), 683-691.
- Erichsen, J.T., Hodos, W., & Evinger, C. (2000). The pupillary light reflex, accommodation and convergence: Comparative considerations. In O. Franzen, H. Richter, & L. Stark (Eds.), *Accommodation and vergence mechanisms in the visual system* (pp. 31-42). Basel: Birkhauser.
- Evans, R.W., & Jacobson, D.M. (2003). Transient anisocoria in a migraineur. *Headache*, 43(4), 416-418. doi:10.1046/j.1526-4610.2003.03081.x
- Fabiani, M., & Donchin, E. (1995). Encoding processes and memory organization: A model of the Von Restorff effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 21(1), 224- 240. doi:10.1037/0278-7393.21.1.224
- Fabiani, M., Gratton, G., & Coles, M.G.H. (2000). Event-related brain potentials. In J.T. Cacioppo, L.G. Tassinary, & G.G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 53-84). New York: Cambridge University Press.
- Fanciullacci, M., Galli, P., Pietrini, U., & Sicuteri, F. (1977). Adrenergic supersensitivity of the pupil in idiopathic headache. *Experientia*, 33(8), 1082–1083. doi:10.1007/BF01945983
- Finnigan, S., Humphreys, M.S., Dennis, S., & Geffen, G. (2002). ERP 'old/new' effects: Memory strength and decisional factor(s). *Neuropsychologia*, 40(13), 2288-2304. doi:10.1016/S0028-3932(02)00113-6
- Fogarty, C., & Stern, J.A. (1989). Eye movements and blinks: Their relationship to higher cognitive processes. *International Journal of Psychophysiology*, 8(1), 35-42. doi:10.1016/0167-8760(89)90017-2
- Forbes, T.R. (1977). Why is it called 'beautiful lady'? A note on belladonna. *Bulletin of the New York Academy of Medicine*, 53(4), 403-406.
- Ford, E.B. (2006). Lie detection: Historical, neuropsychiatric and legal dimensions. *International Journal of Law & Psychiatry*, 29(3), 159-177. doi:10.1016/j.ijlp.2005.07.001
- Foster, M. (1891). *A textbook of physiology volume IV* (5th ed.). Cambridge: C.J. Clay, M.A. & Sons.
- Foster, J.K., & Jellicic, M. (Eds.) (1999). *Memory: Systems, process or function*. Oxford: Oxford University Press.

- Francis, R.D., & Kelly, M.R. (1969). An investigation of the relationship between word stimuli and optical pupil-size. *Australian Journal of Psychology*, 21(2), 117-125. doi:10.1080/00049536908257774
- Frank, M.G., & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stake lies. *Journal of Personality & Social Psychology*, 72(6), 1429-1439. doi:10.1037/0022-3514.72.6.1429
- Franzen, M.D., & Iverson, G.L. (2000). Detecting negative response bias and diagnosing malingering: The dissimulation exam. In J. Snyder, & P.J. Nussbaum (Eds.), *Clinical neuropsychology: A pocket handbook for assessment* (2nd ed.) (pp. 102-119). Washington, DC: American Psychological Association. doi:10.1037/11299-006
- Frensch, P.A. (1998). One concept, multiple meanings: On how to define the concept of implicit learning. In M.A. Stadler, & P.A. Frensch (Eds.), *Handbook of implicit learning* (pp. 47-104). Thousand Oaks, CA: Sage.
- Frensch, P.A., & Miner, C.S. (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory & Cognition*, 22(1), 95-110. doi:10.3758/BF03202765
- Friedman, D. (2004). ERP studies of recognition memory: Differential effects of familiarity, recollection, episodic priming. *Cognitive Sciences*, 1(1), 81-121.
- Friedman, D., Hakerem, G., Sutton, S., & Fleiss, J.L. (1973). Effect of stimulus uncertainty on the pupillary dilation response and the vertex evoked potential. *Electroencephalography & Clinical Neurophysiology*, 34(5), 475-484. doi:10.1016/0013-4694(73)90065-5
- Friedman, D., Hamberger, M., Stern, Y., & Marder, K. (1992). Event-related potentials (ERPs) during repetition priming in Alzheimer's patients and young and older controls. *Journal of Clinical & Experimental Neuropsychology*, 14(4), 448-462. doi:10.1080/01688639208402837
- Friedman, D., & Johnson Jr., R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research & Technique*, 51(1), 6-28. doi:10.1002/1097-0029(20001001)51:1<6::AID-JEMT2>3.0.CO;2-R
- Fukuda, K., Stern, J.A., Brown, T.B., & Russo, M.B. (2005). Cognition, blinks, eye-movements, and pupillary movements during performance of a running memory task. *Aviation, Space, & Environmental Medicine*, 76(1S), C75-C85.

- Gabay, S., Pertzov, Y., & Henik, A. (2011). Orienting of attention, pupil-size, and the norepinephrine system. *Attention, Perception & Psychophysics*, 73(1), 123-129. doi:10.3758/s13414-010-0015-4
- Galappathie, N., & Vakili, K. (2009). Detecting memory malingering: A case study. *British Journal of Forensic Practice*, 11(1), 20-24. doi:10.1108/14636646200900005
- Gambill, H.D., Ogle, K.N., & Kearns, T.P. (1967). Mydriatic effect of four drugs determined with pupillograph. *Archives of Ophthalmology*, 77(6), 740-746.
- Gang, K. (1945). Psychosomatic factors in the control of pupillary movements. *Journal of Clinical Psychopathology Psychotherapy*, 6, 461-472.
- Gao, Y., Barreto, A., & Adjouadi, M. (2010). Affective assessment of a computer user through the processing of the pupil diameter signal. In T. Sobh, & K. Elleithy (Eds.), *Innovations in computing sciences and software engineering* (pp. 189-194). New York: Springer. doi:10.1007/978-90-481-9112-3_32
- Gardiner, J.M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, 16(4), 309-313. doi:10.3758/BF03197041
- Gardiner, J.M., & Java, R.I. (1993). Recognizing and remembering. In S.E. Gathercole, M.A. Conway, E.E. Morris, & A.E Collins (Eds.), *Theories of memory* (pp. 163-188). Hove, England: Erlbaum.
- Gardiner, J.M., & Parkin, A.J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition*, 18(6), 579-583. doi:10.3758/BF03197100
- Gardiner, J.M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving, & F.I.M. Craik (Eds.), *Oxford handbook of memory* (pp. 229-244). New York: Oxford University Press.
- Gardner, R.M., Beltramo, J.S., & Krinsky, R. (1975). Pupillary changes during encoding, storage, and retrieval of information. *Perceptual & Motor Skills*, 41(3), 951-955. doi:10.2466/pms.1975.41.3.951
- Gardner, R.M., Mo, S.S., & Borrego, R. (1974a). Inhibition of pupillary orienting reflex by novelty in conjunction with recognition memory. *Bulletin of the Psychonomic Society*, 3(3-B), 237-238.
- Gardner, R.M., Mo, S.S., & Krinsky, R. (1974b). Inhibition of pupillary orienting reflex by heteromodal novelty. *Bulletin of the Psychonomic Society*, 4(5B), 510-512.

- Gardner, R.M., Philp, P., & Radacy, S. (1978). Pupillary changes during recall in children. *Journal of Experimental Child Psychology*, 25(1), 168-172. doi:10.1016/0022-0965(78)90046-2
- Garrett, J.C., Harrison, D.W., & Kelly, P.L. (1989). Pupillometric assessment of arousal to sexual stimuli: Novelty effects or preference. *Archives of Sexual Behavior*, 18(3), 191-201. doi:10.1007/BF01543194
- Geacintov, T., & Peavler, W.S. (1974). Pupillography in industrial fatigue assessment. *Journal of Applied Psychology*, 59(2), 213-216. doi:10.1037/h0036529
- Ghanem, T., & Early, S.V. (2006). Vagal nerve stimulator implantation: An otolaryngologist's perspective. *Otolaryngology Head Neck Surgery*, 135(1), 46-51. doi:10.1016/j.otohns.2006.02.037
- Ghodse, A.H., Greaves, J.L., & Lynch, D. (1999). Evaluation of the opioid addiction test in an out-patient drug dependency unit. *British Journal of Psychiatry*, 175, 158-162. doi: 10.1192/bjp.175.2.158
- Gillund, G., & Shiffrin, R.M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, 91(1), 1-67. doi:10.1037/0033-295X.91.1.1
- Gilzenrat, M.S. (2006). *The role of the locus coeruleus in cognitive control: Pupillometric measures and computational mechanisms*. (Unpublished doctoral dissertation). Princeton University, New Jersey, USA.
- Gilzenrat, M.S., Cohen, J.D., Rajkowski, J., & Aston-Jones, G. (2003). Pupil dynamics predict changes in task engagement mediated by locus coeruleus. *Society for Neuroscience Abstracts*, 515.19.
- Gilzenrat, M.S., Nieuwenhuis, S., Jepma, M., & Cohen, J.D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, 10(2), 252-269. doi:10.3758/CABN.10.2.252
- Glanzer, M., & Adams, J.K. (1985). The mirror effect in recognition memory. *Memory & Cognition*, 13(1), 8-20.
- Goldmann, R.E., Sullivan, A.L., Droller, D.B., Rugg, M.D., Curran, T., Holcomb, P.J., ... Budson, A.E. (2003). Late frontal brain potentials distinguish true and false recognition. *Neuroreport*, 14(13), 1717-1720. doi:10.1097/01.wnr.0000087908.78892.23
- Goldwater, B. (1972). Psychological significance of pupillary movements. *Psychological Bulletin*, 77(5), 340-355. doi:10.1037/h0032456

- Goodrich, G.L. (1974). *Actual and apparent pupillary diameters by sex, age, refraction, iris coloration, and fundus coloration*. (Unpublished doctoral dissertation). Washington State University, USA.
- Gorman, A.M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *Journal of Experimental Psychology*, 61(1), 23-29. doi:10.1037/h0040561
- Gouvier, W.D., Prestholdt, P.H., & Warner, M.S. (1988). A survey of common misconceptions about head injury and recovery. *Archives of Clinical Neuropsychology*, 3(4), 331-343. doi:10.1093/arclin/3.4.331
- Graf, P., Mandler, G., & Haden, P.E. (1982). Simulating amnesic symptoms in normal subjects. *Science*, 218(4578), 1243-1244. doi:10.1126/science.7146909
- Granholm, E., Asarnow, R.F., Sarkin, A.J., & Dykes, K.L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, 33(4), 457-461. doi:10.1111/j.1469-8986.1996.tb01071.x
- Granholm, E., Morris, S., Galasko, D., Shults, C., Rogers, E., & Vukov, B. (2003). Tropicamide effects on pupil-size and pupillary light reflexes in Alzheimer's and Parkinson's disease. *International Journal of Psychophysiology*, 47(2), 95-115. doi:10.1016/S0167-8760(02)00122-8
- Granholm, E., Morris, S.K., Sarkin, A.J., Asarnow, R.F., & Jeste, D.V. (1997). Pupillary responses index overload of working memory resources in schizophrenia. *Journal of Abnormal Psychology*, 106(3), 458-467. doi:10.1037/0021-843X.106.3.458
- Granholm, E., & Verney, S.P. (2004). Pupillary responses and attentional allocation on the visual backward masking task in schizophrenia. *International Journal of Psychophysiology*, 52, 37-52. doi:10.1016/j.ijpsycho.2003.12.004
- Gratiolet, P., & Grandeau, L. (1865). *De la physionomie et des mouvements d'expression: Suivi d'une notice sur sa vie et ses travaux et de la nomenclature de ses ouvrages*. Paris: J. Hetzel.
- Green, P., & Astner, K. (1995). *Manual for the Oral Word Memory Test*. Durham, NC: CogniSyst, Inc.
- Greiffenstein, M.F., Baker, J.W., Gola, T., Donders, J., & Miller, L. (2002). The fake bad scale in atypical and severe closed head injury litigants. *Journal of Clinical Psychology*, 58(12), 1591-1600. doi:10.1002/jclp.10077

- Gregg, V.H. (1976). Word frequency, recognition, and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183-216). London: John Wiley & Sons.
- Greve, A., Donaldson, D.I., & van Rossum, M.C.W. (2010). A single-trace dual-process model of episodic memory: A novel computational account of familiarity and recollection. *Hippocampus*, 20(2), 235-51.
doi:10.1002/hipo.20606
- Grill-Spector, K., & Malach, R. (2001). fMRI-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, 107(1-3), 293-321.
- Gronlund, S.D., & Ratcliff, R. (1989). Time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 15(5), 846-858.
doi:10.1037/0278-7393.15.5.846
- Groppe, D.M., Urbach, T.P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, 48(12), 1711-1725. doi:10.1111/j.1469-8986.2011.01273.x
- Groves, D.A., & Brown, V.J. (2005). Vagus nerve stimulation: A review of its applications and potential mechanisms that mediate its clinical effects. *Neuroscience & Biobehavioral Reviews*, 29(3), 493-500.
doi:10.1016/j.neubiorev.2005.01.004
- Gruber, T., Tsivilis, D., Giabbiconi, C.M., & Muller, M.M. (2008). Induced electroencephalogram oscillations during source memory: Familiarity is reflected in the gamma band, recollection in the theta band. *Journal of Cognitive Neuroscience*, 20(6), 1043-1053. doi:10.1162/jocn.2008.20068
- Grunberger, J., Linzmayer, L., Majda, E.M., Reitner, A., & Walter, H. (1996). Pupillary dilatation test and Fourier analysis of pupillary oscillations in patients with multiple sclerosis. *European Archives of Psychiatry & Clinical Neuroscience*, 246(4), 209-212. doi:10.1007/BF02188955
- Grunberger, J., Linzmayer, L., Walter, H., Stohr, H., Saletu-Zyhlarz, G.M., Grunberger, M., & Lesch, O.M. (1998). Psychophysiological diagnostics in alcohol dependency: Fourier analysis of pupillary oscillations and the receptor test for determination of cholinergic deficiency. *Alcohol & Alcoholism*, 33(5), 541-548.

- Guinan, J.F. (1967). *An investigation of the relationship between pupil-size and emotional words*. (Doctoral dissertation). Dissertation Abstracts, 27(9-BO), 3286-3287.
- Guo, C., Lawson, A.L., & Jiang, Y. (2007). Distinct neural mechanisms for repetition effects of visual objects. *Neuroscience*, 149(4), 747-759.
doi:10.1016/j.neuroscience.2007.07.060
- Habib, R., & LePage, M. (1999). Novelty assessment in the brain. In E. Tulving (Ed.), *Memory, consciousness, and the brain* (pp. 265-277). Philadelphia, PA: Psychology Press..
- Hack Tuke, D. (1884). *Illustrations of the Influence of the mind upon the body in health and disease: Designed to elucidate the action of the imagination* (2nd ed.). Philadelphia, PA: Henry C Lea's Son & Co.
- Hakerem, G. (1973). The effect of cognitive manipulation on pupillary diameter and evoked vertex potentials. In E. Dodt, & K.E. Schrader (Eds.), *The Normal and The Perturbed Motion Pupil Symposium of the The German Society of Ophthalmology*. Munich: J.F. Bergmann Verlag.
- Hakerem, G. (1974). Conceptual stimuli, pupillary dilation and evoked cortical potentials. In M.P. Janisse (Ed.), *Pupillary dynamics and behavior* (pp. 135-158). Plenum Press, New York.
- Hakerem, G., & Sutton, S. (1966). Pupillary response at visual threshold. *Nature*, 212, 485-486. doi:10.1038/212485a0
- Hannula, D.E., & Ranganath, C. (2009). The eyes have it: Hippocampal activity predicts expression of memory in eye movements. *Neuron*, 63(5), 592-599.
doi:10.1016/j.neuron.2009.08.025
- Hansen, D.W., & Pece, A.E. (2005). Eye tracking in the wild. *Computer Vision and Image Understanding*, 98(1), 155-181. doi:10.1016/j.cviu.2004.07.013
- Harle, D.E., Wolffsohn, J.S., & Evans, B.J.W. (2005). The pupillary light reflex in migraine. *Ophthalmic & Physiological Optics*, 25(3), 240-245.
doi:10.1111/j.1475-1313.2005.00291.x
- Harvey, W. (1649). *Exercitatio duae anatomica de circulatione sanguinis ad Joannem Riolanum filium*. Cambridge: Ex officina Rogeri Danielis.
- Hawkes, C.H., & Stow, B. (1981). Pupil-size and the pattern evoked visual response. *Journal of Neurology, Neurosurgery, and Psychiatry*, 44(1), 90-91.
doi:10.1136/jnnp.44.1.90

- Heaver, B., & Hutton, S.B. (2010). Keeping an eye on the truth: Pupil-size, recognition memory and malingering. *International Journal of Psychophysiology*, 77(3), 306-306. doi:10.1016/j.ijpsycho.2010.06.206
- Heaver, B., & Hutton, S.B. (2011). Keeping an eye on the truth? Pupil-size changes associated with recognition memory. *Memory*, 19(4), 398-405. doi:10.1080/09658211.2011.575788
- Heilbrun, K., Bennett, W.S., White, A.J., & Kelly, J. (1990). An MMPI-based empirical model of malingering and deception. *Behavioral Sciences & the Law*, 8(1), 45-53. doi:10.1002/bsl.2370080106
- Heilveil, I. (1976). Deception and pupil-size. *Journal of Clinical Psychology*, 32(3), 675-676.
- Heinrich, W. (1896). Die aufmerksamkeit und die funktion der sinnesorgane [The attention and the function of the sensory organs]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 9, 342-388.
- Helmholtz, H. (1853). Ueber einige gesetze der vertheilung elektrischer ströme in körperlichen leitern mit anwendung auf die theirisch-elektrischen versuche [Some laws about the distribution of electrical currents in volume conductors, with application to animal electric experiments]. *Annalen der Physik und Chemie*, 89, 211-233.
- Henson, R.N., Hornberger, M., & Rugg, M.D. (2005). Further dissociating processes in recognition memory using fMRI. *Journal of Cognitive Neuroscience*, 17(7), 1058-1073. doi:10.1162/0898929054475208
- Henson, R.N.A., Rugg, M.D., Shallice, T., Josephs, O., & Dolan, R.J. (1999). Recollection and familiarity in recognition memory: An event-related functional magnetic resonance imaging study. *Journal of Neuroscience*, 19(10), 3962-3972.
- Henson, R.N.A., Shallice, T., Gorno-Tempini, M.L., & Dolan, R.J. (2002). Face repetition effects in implicit and explicit memory tests as measured by fMRI. *Cerebral Cortex*, 12(2), 178-186. doi:10.1093/cercor/12.2.178
- Hertel, P.T., & Parks, C. (2002). Emotional episodes facilitate word recall. *Cognition & Emotion*, 16(5), 685-694. doi:10.1080/02699930143000491
- Hess, E.H. (1965). Attitude and pupil-size. *Scientific American*, 212, 46-54.
- Hess, E.H. (1972). Pupillometrics. In N.S. Greenfield, & R.A. Sternback (Eds.), *Handbook of psychophysiology* (pp. 491-531). New York: Holt, Rinehart & Winston.

- Hess, E.H. (1975). *The tell-tale eye*. New York: Van Nostrand Reinhold.
- Hess, E.H., Beaver, P.W., & Shrout, P.E. (1975). Brightness contrast effects in a pupillometric experiment. *Perception & Psychophysics*, 18(2), 125-127. doi:10.3758/BF03204099
- Hess, E.H., & Polt, J.M. (1960). Pupil-size as related to interest value of visual stimuli. *Science*, 132(3423), 349-350. doi:10.1126/science.132.3423.349.
- Hess, E.H., & Polt, J.M. (1964). Pupil-size in relation to mental activity in simple problem solving. *Science*, 143(3611), 1190-1192. doi:10.1126/science.143.3611.1190
- Hess, E.H., Seltzer, A.L., & Shlien, J.M. (1965). Pupil response of hetero- and homosexual males to pictures of men and women: A pilot study. *Journal of Abnormal Psychology*, 70(3), 165-168. doi:10.1037/h0021978
- Hicks, R.A., Reaney, T., & Hill, L. (1967). Effects of pupil-size and facial angle on preference for photographs of a young woman. *Perceptual & Motor Skills*, 24(2), 388-390. doi:10.2466/pms.1967.24.2.388
- Hillyard, S.A., Hink, R.F., Schwent, V.L., & Picton, T.W. (1973). Electrical signs of selective attention in the human brain. *Science*, 182(4108), 177-180. doi:10.1126/science.182.4108.177
- Hink, R.F., Van Voorhis, S.T., Hillyard, S.A., & Smith, T.S. (1977). The division of attention and the human auditory evoked potential. *Neuropsychologia*, 15(4-5), 597-605. doi:10.1016/0028-3932(77)90065-3
- Hintzman, D.L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95(4), 528-551.
- Hintzman, D.L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory & Language*, 33, 1-18.
- Hintzman, D.L., & Curran, T. (1995). When encoding fails: Instructions, feedback, and registration without learning. *Memory & Cognition*, 23(2), 213-226. doi:10.3758/BF03197223
- Hintzman, D.L., & Curran, T. (1997). Comparing retrieval dynamics in recognition memory and lexical decision. *Journal of Experimental Psychology: General*, 126(3), 228-247.
- Hintzman, D.L., Curran, T., & Oppy, B. (1992). Effects of similarity and repetition on memory: Registration without learning? *Journal of Experimental Psychology:*

- Learning, Memory & Cognition*, 18(4), 667-680. doi:10.1037/0278-7393.18.4.667
- Holdstock, J.S., Mayes, A.R., Roberts, N., Cezayirli, E., Isaac, C.L., O'Reilly, R.C., & Norman, K.A. (2002). Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus*, 12, 341-351. doi:10.1002/hipo.10011
- Hou, R.H., Freeman, C., Langley, R.W., Szabadi, E., & Bradshaw, C.M. (2005). Does modafinil activate the locus coeruleus in man? Comparison of modafinil and clonidine on arousal and autonomic functions in human volunteers. *Psychopharmacology*, 181(3), 537-549. doi:10.1007/s00213-005-0013-8
- Howard, D.V., & Howard, J.H. (1989). Age differences in learning serial patterns: Direct versus indirect measures. *Psychology & Ageing*, 4(3), 357-364. doi:10.1037/0882-7974.4.3.357
- Howard, D.V., & Howard, J.H. (1992). Adult age differences in the rate of learning serial patterns: Evidence from direct and indirect tests. *Psychology & Aging*, 7(2), 232-241. doi:10.1037/0882-7974.7.2.232
- Huber, D.E., Clark, T.F., Curran, T., & Winkielman, P. (2008). Effects of repetition priming on recognition memory: Testing a perceptual fluency-disfluency model. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 34(6), 1305-1324. doi:10.1037/a0013370
- Huettel, S.A., Song, A.W., & McCarthy, G. (2004). *Functional magnetic resonance imaging*. Massachusetts, USA: Sinauer Associates Inc
- Humphreys, M.S., Bain, J.D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review*, 96(2), 208-233.
- Hupé, J.M., Lamirel, C., & Lorenceau, J. (2009). Pupil dynamics during bistable motion perception. *Journal of Vision*, 9(7), 10. doi:10.1167/9.7.10
- Hutchinson, G.L. (2001). *Disorders of simulation: Malingering, factitious disorders, and compensation neurosis*. Madison, Connecticut: Psychosocial Press.
- Hyönä, J., Tammola, J., & Alaja, A.M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology: A*, 48(3), 598-612. doi:10.1080/14640749508401407
- Iijima, A., Haida, M., Ishikawa, N., Ueno, A., Minamitani, H., & Shinohara, Y. (2003). Re-evaluation of tropicamide in the pupillary response test for Alzheimer's

- disease. *Neurobiology of Aging*, 24(6), 789-796. doi:10.1016/S0197-4580(02)00235-X
- Ingram, K.M., Mickes, L., & Wixted, J.T. (2011). Recollection can be weak and familiarity can be strong. *Journal of Experimental Psychology. Learning, Memory, & Cognition*, 38(2), 325-339. doi:10.1037/a0025483
- Iqbal, S.T., Adamczyk, P.D., Zheng, X.S., & Bailey, B.P. (2005). Towards an index of opportunity: Understanding changes in mental workload during task execution. In Proceedings from *The SIGCHI Conference on Human Factors in Computing Systems* (pp. 311-320). Portland, Oregon: ACM Press.
- Iqbal, S.T., Zheng, X.S., & Bailey, B.P. (2004). Task-evoked pupillary response to mental workload in human-computer interaction. In Proceedings from *The SIGCHI Conference on Human Factors in Computing Systems* (pp. 1477-1480). New York: ACM Press. doi:10.1145/985921.986094
- Isidore of Seville (2006). *The etymologies of Isidore of Seville* (S.A. Barney, W.J. Lewis, J.A. Beach, & O. Berghof, Trans.). Cambridge: Cambridge University Press. (Original work published 636 AD).
- Jacoby, L.L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology-General*, 110(3), 306-340. doi:10.1037/0096-3445.110.3.306
- Jacoby, L.L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, 118(2), 126-135. doi:10.1037/0096-3445.118.2.126
- Jain, J., Siegle, G.J., Gu, C., Moore, C.G., Ivanco, L.S., Studenski, S., ... Steinhauer, S.R. (2011). Pupillary unrest correlates with arousal symptoms and motor signs in Parkinson disease. *Movement Disorders*, 26(7), 1344-1337. doi:10.1002/mds.23628
- Jainta, S., & Baccino, T. (2010). Analyzing the pupil response due to increased cognitive demand: An independent component analysis study. *International Journal of Psychophysiology*, 77, 1-7. doi:10.1016/j.ijpsycho.2010.03.008
- Jainta, S., Hoormann, J., & Jaschinski, W. (2008). Ocular accommodation and cognitive demand: An additional indicator besides pupil-size and cardiovascular measures? *Journal of Negative Results in BioMedicine*, 7(6). doi:10.1186/1477-5751-7-6
- James, W. (1890). *The principles of psychology, vol. 1*. New York: Henry Holt & Co.

- Janisse, M.P. (1973). Pupil-size and affect: A critical review of the literature since 1960. *Canadian Psychologist*, 14(4), 311-329. doi:10.1037/h0082230
- Janisse, M.P. (Ed.) (1974). *Pupillary dynamics and behavior*. New York, NY: Plenum Press.
- Janisse, M.P. (1976). The relationship between pupil-size and anxiety: A review. In I.G. Sarason, & C.D. Spielberger (Eds.), *Stress and anxiety* (Vol 3). Oxford, England: Hemisphere.
- Janisse, M.P. (1977). *Pupillometry: The psychology of the pupillary response*. Washington, DC: Hemisphere Publishing Corporation.
- Janisse, M.P., & Bradley, M.T. (1980). Deception, information and the pupillary response. *Perceptual & Motor Skills*, 50(3), 748-750. doi: 10.2466/pms.1980.50.3.748
- Janisse, M.P., & Peavler, W.S. (1974). Pupillary response today: Emotion in the eye. *Psychology Today*, 7, 60-63.
- Jasper, H.A. (1958). The ten–twenty electrode system of the International Federation. *Electroencephalography & Clinical Neurophysiology*, 10(2), 371–375. doi:10.1016/0013-4694(58)90053-1
- Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation tradeoff: Evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, 23(7), 1587-1596. doi:10.1162/jocn.2010.21548
- Jiang, Y., Haxby, J.V., Martin, A., Ungerleider, L.G., & Parasuraman, R. (2000). Complementary neural mechanisms for tracking items in human working memory. *Science*, 287(5453), 643-646. doi:10.1126/science.287.5453.643
- Jin, P. (1992). Towards a reconceptualization of the law of initial value. *Psychological Bulletin*, 111(1), 176-184. doi:10.1037/0033-2909.111.1.176
- Johnson, D.A. (1971). Pupillary responses during a short-term memory task: Cognitive processing, arousal, or both? *Journal of Experimental Psychology*, 90(2), 311-318. doi:10.1037/h0031562
- Johnson, R.J. (1986). A triarchic model of P300 amplitude. *Psychophysiology*, 23(4), 367-384. doi:10.1111/j.1469-8986.1986.tb00649.x
- Johnson, R.J. (1995). Event-related potential insights into the neurobiology of memory systems. In F. Boller, & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 10, pp. 135-163). Amsterdam: Elsevier.

- Jones, T.C., & Jacoby, L.L. (2005). Conjunction errors in recognition memory: Modality-free errors for older adults but not young adults. *Acta Psychologica*, 120(1), 55-73. doi:10.1016/j.actpsy.2005.03.003
- Jouvet, M. (1969). Biogenic amines and the states of sleep. *Science*, 163(3862), 32-41. doi:10.1126/science.163.3862.32
- Just, M.A., & Carpenter, P.A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, 47(2), 310-339.
- Just, M.A., & Carpenter, P.A. (1995). Spatial working memory: Cognitive capacity in mental rotation. Unpublished manuscript, Carnegie Mellon University, Pittsburgh, PA.
- Just, M.A., Carpenter, P.A., Keller, T.A., Eddy, W.F., & Thulborn, K.R. (1996). Brain activation modulated by sentence comprehension. *Science*, 274(5284), 114-116. doi:10.1126/science.274.5284.114
- Just, M.A., Carpenter, P.A., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science*, 4(1-2), 56-88. doi:10.1080/14639220210159735
- Kafkas, A., & Montaldi, D. (2011). Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection. *The Quarterly Journal of Experimental Psychology*, 64(10), 1971-1989. doi:10.1080/17470218.2011.588335
- Kahneman, D. (1971). Changes in pupil-size and visual discrimination during mental effort. In J.R. Pierce, & J.R. Levene (Eds.), *Visual science: Proceedings of the 1968 International Symposium*. Bloomington, IN: Indiana University Press.
- Kahneman, D. (1973). *Attention and effort*. New Jersey: Prentice Hall.
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, 154(3756), 1583-1585. doi:10.1126/science.154.3756.1583
- Kahneman, D., Beatty, J., & Pollack, I. (1967). Perceptual deficit during a mental task. *Science*, 157(3785), 218-219. doi:10.1126/science.157.3785.218
- Kahneman, D., Onuska, L., & Wolman, R.E. (1968a). Effects of grouping on the pupillary response in a short-term memory task. *Quarterly Journal of Experimental Psychology*, 20(3), 309-311. doi:10.1080/14640746808400168

- Kahneman, D., & Peavler, W.S. (1969). Incentive effects and pupillary changes in association learning. *Journal of Experimental Psychology*, 79(2), 312-318. doi:10.1037/h0026912
- Kahneman, D., Peavler, W.S., & Onuska, L. (1968b). Effects of verbalization and incentive on the pupil response to mental activity. *Canadian Journal of Psychology*, 22(3), 186-196. doi:10.1037/h0082759
- Kahneman, D., Tursky, B., Shapiro, D., & Crider, A. (1969). Pupillary, heart rate, and skin resistance changes during a mental task. *Journal of Experimental Psychology*, 79(1), 164-167. doi:10.1037/h0026952
- Kahneman, D., & Wright, P. (1971). Changes of pupil-size and rehearsal strategies in a short-term memory task. *Quarterly Journal of Experimental Psychology*, 23(2), 187-196. doi:10.1080/14640747108400239
- Kang, M.J., Hsu, M., Krajbich, I.M., Loewenstein, G., McClure, S.M., Wang, J.T., & Camerer, C.F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, 20(8), 963-973. doi:10.1111/j.1467-9280.2009.02402.x
- Kapucu, A., Macmillan, N.A., & Rotello, C.M. (2010). Positive and negative remember judgments and ROCs in the plurals paradigm: Evidence for alternative decision strategies. *Memory & Cognition*, 38(5), 541-554. doi:10.3758/MC.38.5.541
- Kardon, R. (1995). Pupillary light reflex. *Current Opinion in Ophthalmology*, 6(6), 20-26.
- Karis, D., Fabiani, M., & Donchin, E. (1984). "P300" and memory: Individual differences in the von Restorff effect. *Cognitive Psychology*, 16(2), 177-216. doi:10.1016/0010-0285(84)90007-0
- Kasikci, T., & Bek, S. (2010). Malingering in neurology clinics. *Turkiye Klinikleri Journal of Medical Sciences*, 30(4S), S35-S42.
- Kelley, R., & Wixted, J.T. (2001). On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 27(3), 701-722. doi:10.1037/0278-7393.27.3.701
- Kirschen, M.P., Davis-Ratner, M.S., Jerde, T.E., Schraedley-Desmond, P., & Desmond, J.E. (2006). Enhancement of phonological memory following transcranial magnetic stimulation. *Behavioural Neurology*, 17, 187-194.

- Kiebel, S.J., & Friston, K.J. (2004). Statistical parametric mapping for event-related potentials: I. Generic considerations. *NeuroImage*, 22(2), 492-502. doi:10.1016/j.neuroimage.2004.02.012
- Kim, M., Beversdorf, D.Q., & Heilman, K.M. (2000). Arousal response with aging: Pupillographic study. *Journal of the International Neuropsychological Society*, 6(3), 348-350. doi:10.1017/S135561770000309X
- Kintsch, W. (1967). Memory and decision aspects of recognition learning. *Psychological Review*, 74(6), 496-504. doi:10.1037/h0025127
- Klimesch, W. (1995). Memory processes described as brain oscillations in the EEG-alpha and theta bands. *Psychology*, 95(6), 23-37.
- Knaggs, R.D., Crighton, I.M., Cobby, T.F., Fletcher, A.J.P., & Hobbs, G.J. (2004). The pupillary effects of intravenous morphine, codeine, and tramadol in volunteers. *Anesthesia & Analgesia*, 99(1), 108-112. doi:10.1213/01.ANE.0000116924.16535.BA
- Knopman, D., & Nissen, M.I. (1987). Implicit learning in patients with probable Alzheimer's disease. *Neurology*, 37(5), 784-788.
- Knowlton, B.J., Ramus, S., & Squire, L.R. (1992). Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science*, 3(3), 172-179. doi:10.1111/j.1467-9280.1992.tb00021.x
- Kohn, M., & Clynes, M. (1969). Color dynamics of the pupil. *Annals of the New York Academy of Sciences*, 156, 931-950. doi:10.1111/j.1749-6632.1969.tb14024.x
- Koikegami, H., & Yoshida, K. (1953). Pupillary dilatation induced by stimulation of amygdaloid nuclei. *Psychiatry & Clinical Neurosciences*, 7(2), 109-126. doi:10.1111/j.1440-1819.1953.tb00600.x
- Kok, A. (1997). Event-related potential (ERP) reflections of mental resources: A review and synthesis, *Biological Psychology*, 45(1-3), 19-56. doi:10.1016/S0301-0511(96)05221-0
- Konishi, S., Wheeler, M.E., Donaldson, D.I., & Buckner, R.L. (2000). Neural correlates of retrieval success. *Neuroimage*, 12(3), 276-286. doi:10.1006/nimg.2000.0614
- Kramer, A.F. (1991). Physiological metrics of mental workload: A review of recent progress. In D.I. Damos (Ed.), *Multi-task performance* (pp. 279-323). London: Taylor & Francis.

- Krantz, D.H. (1969). Threshold theories of signal detection. *Psychological Review*, 76(3), 308-324. doi:10.1037/h0027238
- Krinsky, R., & Gardner, R.M. (1977). Effects of retention interval on pupillary responses during a simple discrimination. *Perceptual & Motor Skills*, 45(2), 457-458. doi:10.2466/pms.1977.45.2.457
- Kruger, P.B. (1975). Luminance changes of the fundus reflex. *American Journal of Optometry & Physiological Optics*, 52(12), 847-861.
- Kruger, P.B. (1977a). Changes in fundus reflex luminance with increased cognitive processing. *American Journal of Optometry & Physiological Optics*, 54(7), 445-51.
- Kruger, P.B. (1977b). The role of accommodation in increasing the luminance of the fundus reflex during cognitive processing. *Journal of the American Optometric Association*, 48(12), 1493–1496.
- Kucera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Kuipers, J.R., & Thierry, G. (2011). N400 amplitude reduction correlates with an increase in pupil-size. *Frontiers in Human Neuroscience*, 5(61), 1-5. doi:10.3389/fnhum.2011.00061
- Kumnick, L.S. (1954). Pupillary psychosensory restitution and aging. *Journal of the Optical Society of America*, 44(9), 735-741. doi:10.1364/JOSA.44.000735
- Kumnick, L.S. (1956). Aging and pupillary response to light and sound stimuli. *Journal of Gerontology*, 11(1), 38-45. doi:10.1093/geronj/11.1.38
- Kuntz, A. (1929). *The autonomic nervous system*. Philadelphia: Lea & Febiger.
- Kutas, M., & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205. doi:10.1126/science.7350657
- Lacey, J.I. (1959). Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome. In E.A. Rubinstein, & M.B. Parloff (Eds.), *Research in psychotherapy* (pp. 160-208). Washington, DC: American Psychological Association.
- Lacey, J.I. (1967). Somatic response patterning and stress: Some revisions of activation theory. In M.H. Appley, & R. Trumbull (Eds.), *Psychological stress: Issues for research* (pp. 14-42). New York: Appleton-Century-Crofts.
- Laeng, B., Waterloo, K., Johnsen, S.H., Bakke, S.J., Låg, T., Simonsen, S.S., & Høgsæt, J. (2007). The eyes remember it: Oculography and pupillometry

- during recollection in three amnesic patients. *Journal of Cognitive Neuroscience*, 19(11), 1888-1904. doi:10.1162/jocn.2007.19.11.1888
- Lanatà, A., Armato, A., Valenza, G., & Scilingo, E.P. (2011). Eye tracking and pupil-size variation as response to affective stimuli: A preliminary study. In Proceedings from *The Pervasive Computing Technologies for Healthcare 5th International Conference*. Dublin, Ireland: Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. doi:10.4108/icst.pervasivehealth.2011.246056
- Lange, G., Janal, M.N., Maniker, A., FitzGibbons, J., Fobler, M., Cook, D., & Natelson, B.H. (2011). Safety and efficacy of vagus nerve stimulation in fibromyalgia: A phase I/II proof of concept trial. *Pain Medicine*, 12(9), 1406-1413. doi:10.1111/j.1526-4637.2011.01203.x
- Lange, R.T., Senior, G., Douglas, L., & Dawes, S. (2003). Base rates and clinical utility of the WMS-III Rarely Missed Index: Identifying cognitive exaggeration in a head injury litigant sample. *Archives of Clinical Neuropsychology*, 18, 786.
- Lange, R.T., Sullivan, K., & Anderson, D. (2005). Ecological validity of the WMS-III Rarely Missed Index in personal litigation. *Journal of Clinical & Experimental Neuropsychology*, 27(4), 412-424. doi:10.1080/13803390490520319
- Langeluddecke, P.M. (2004). Validation of the Rarely Missed Index (RMI) in detecting memory malingering in mild head injury litigants. *Journal of Forensic Neuropsychology*, 4(1), 49-64.
- Larrabee, G.J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist*, 17(3), 410-425. doi:10.1076/clin.17.3.410.18089
- Larsen, R.J., Mercer, K.A., & Balota, D.A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion*, 6(1), 62-72. doi:10.1037/1528-3542.6.1.62
- Lawless, J.C., & Wake, F.R. (1969). Sex differences in pupillary response to visual stimuli. *Psychophysiology*, 5(5), 568.
- Leal, G., Neves, C., & Vieira, P. (2011). Wavelet analysis of the autonomic outflow of the pupil's variation in normal subjects, given stimuli. In Proceedings from *The 1st Portuguese Meeting in Biomedical Engineering* (pp. 1-4). Lisbon, Portugal: IEEE. doi:10.1109/ENBENG.2011.6026066

- Leding, J.K., & Lampinen, J.M. (2009). Memory conjunction errors: The effects of presentation duration and study repetition. *Memory*, 17(5), 597-607. doi:10.1080/09658210902984518
- Ledoux, K., Pickett, E., Van Droof, L.V., Buz, E., Billings, N.M., & Gordon, B. (2010). Receptive vocabulary knowledge in individuals with autism as assessed by eye movements, pupillary dilation, and event-related potentials. Presented at *The 2010 Neurobiology of Autism Conference*. San Diego.
- Lees-Haley, P.R., English, L.T., & Glenn, W.J. (1991). A fake bad scale for the MMPI-2 for personal injury claimants. *Psychological Reports*, 68, 203-210.
- Lehr, D.J., & Bergum, B. (1966). Note on pupillary adaptation. *Perceptual & Motor Skills*, 23(3), 917-918. doi:10.2466/pms.1966.23.3.917
- Liakos, A., & Crisp, A.H. (1971). Pupil-size in psychoneurotic patients. *Psychotherapy & Psychosomatics*, 19(1-2), 104-110. doi:10.1159/000286311
- Libby, W.L., Lacey, B.C., & Lacey, J.I. (1973). Pupillary and cardiac activity during visual attention. *Psychophysiology*, 10(3), 270-294. doi:10.1111/j.1469-8986.1973.tb00526.x
- Lin, T., Imamiya, A., & Mao, X. (2008). Using multiple data sources to get closer insights into user cost and task performance. *Interacting with Computers*, 20(3), 364-374. doi:10.1016/j.intcom.2007.12.002
- Lockhart, R.S., & Murdock, B.B. (1970). Memory and the theory of signal detection, *Psychological Review*, 74(2), 100-109. doi:10.1037/h0029536
- Loewenfeld, I. (1999). *The pupil: Anatomy, physiology, and clinical applications* (2nd ed., Vol. 1). Oxford, UK: Butterworth-Heinemann.
- Loewenfeld, I.E. (1958). Mechanisms of reflex dilatation of the pupil; historical review and experimental analysis. *Documenta Ophthalmologica Proceedings Series*, 12, 185-448. doi:10.1007/BF00913471
- Loewenfeld, I.E. (1966). Comment on Hess' findings. *Survey of Ophthalmology*, 11, 291-294.
- Loewy, A.D. (1990). Autonomic control of the eye. In A.D. Loewy, & K.M. Spyer (Eds.), *Central regulation of autonomic function* (pp. 268-285). New York: Oxford University Press.
- Lowenstein, O. (1955). Pupillary reflex shapes and topical clinical diagnosis. *Neurology*, 5(9), 631-644.

- Löwenstein, O., & Friedman, E.D. (1942). Pupillographic studies. 1. Present state of pupillography: Its method and diagnostic significance. *Archives of Ophthalmology*, 27(5), 969-993.
- Löwenstein, O., & Loewenfeld, I.E. (1952). Disintegration of central autonomic regulation during fatigue. *Journal of Nervous and Mental Disease*, 115(2), 121-145.
- Löwenstein, O., & Loewenfeld, I.E. (1958). Electronic pupillography: A new instrument and some clinical applications. *American Medical Association Archives of Ophthalmology*, 59(3), 352-363.
- Löwenstein, O., & Loewenfeld, I.E. (1962). The pupil. In H. Davson (Ed.), *The eye: Muscular mechanisms* (pp. 301-340). New York: Academic Press.
- Löwenstein, O., & Loewenfeld, I.E. (1964). The sleep-waking cycle and pupillary activity. *Annals of the New York Academy of Sciences*, 117, 142-156. doi:10.1111/j.1749-6632.1964.tb48169.x
- Lubow, R.E., & Fein, O. (1996). Pupillary size in response to a visual guilty knowledge test: New technique for the detection of deception. *Journal of Experimental Psychology-Applied*, 2(2), 164-177. doi:10.1037//1076-898X.2.2.164
- Luck, S.J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: The MIT Press.
- Luck, S.J. (2010). Is it legitimate to compare conditions with different numbers of trials? Retrieved 17 November 2011 from http://erpinfo.org/Members/sjluck/Mean_Peak_Noise.pdf/view
- Luu, P., & Ferree, T. (2000). Determination of the Geodesic Sensor Nets' electrode positions and their 10-10 international equivalents: Technical note. Eugene, OR: Electrical Geodesics, Inc. Retrieved 26 June 2011 from ftp://ftp.egi.com/pub/documentation/technotes/200_ElectrodePositions.pdf
- Lykken, D.T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology*, 43(6), 385-388. doi:10.1037/h0046060
- Lykken, D.T. (1960). The validity of the guilty knowledge technique: The effects of faking. *Journal of Applied Psychology*, 44(4), 258-262. doi:10.1037/h0044413
- Lykken, D.T. (1972). Range correction applied to heart rate and GSR data. *Psychophysiology*, 9(3), 373-379. doi:10.1111/j.1469-8986.1972.tb03222.x
- Lykken, D.T. (1998). *A tremor in the blood: Uses and abuses of the lie detector*. New York: Plenum Trade.

- Lynn, R. (1966). *Attention, arousal, and the orientation reaction*. Oxford: Pergamon Press.
- Machizawa, M.G., Kalla, R., Walsh, V., & Otten, L.J. (2010). The time course of ventrolateral prefrontal cortex involvement in memory formation. *Journal of Neurophysiology*, 103(3), 1569-1579. doi:10.1152/jn.90937.2008
- MacKenzie, G., & Donaldson, D.I. (2007). Dissociating recollection from familiarity: Electrophysiological evidence that familiarity for faces is associated with a posterior old/new effect. *NeuroImage*, 36, 454-463. doi:10.1016/j.neuroimage.2006.12.005
- MacLeod, C.M., & Kampe, K.E. (1996). Word frequency effects on recall, recognition, and word fragment completion tests. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22(1), 132-142. doi:10.1037/0278-7393.22.1.132
- Macmillan, N.A., & Creelman, C.D. (1990). Response bias: Characteristics of detection theory, threshold theory, and "nonparametric" indexes. *Psychological Bulletin*, 107(3), 401-413. doi:10.1037//0033-2909.107.3.401
- Macmillan, N.A., & Rotello, C.M. (2006). Deciding about decision models of remember and know judgments: A reply to Murdock. *Psychological Review*, 113, 657-665. doi:10.1037/0033-295X.113.3.657
- Malmberg, K.J. (2008). Recognition memory: A review of the critical findings and an integrated theory for relating them. *Cognitive Psychology*, 57(4), 335-384. doi:10.1016/j.cogpsych.2008.02.004
- Malmberg, K.J. (2010). An annotated reading list for recognition memory. Retrieved 14 October 2011 from <http://uweb.cas.usf.edu/~malmberg/Recognition%20Readings.htm>
- Malmberg, K.J., Holden, J.E., & Shiffrin, R.M. (2004). Modeling the effects of repetitions, similarity, and normative word frequency on judgments of frequency and recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 30(2), 319-331. doi:10.1037/0278-7393.30.2.319
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, 87(3), 252-271. doi:10.1037/0033-295X.87.3.252
- Mandler, G. (2008). Familiarity breeds attempts: A critical review of dual process theories of recognition. *Perspectives in Psychological Science*, 3(5), 392-401. doi:10.1111/j.1745-6924.2008.00087.x

- Mandler, G., & Boeck, W. (1974). Retrieval processes in recognition. *Memory & Cognition*, 2(4), 613-615. doi:10.3758/BF03198129
- Manns, J.R., Hopkins, R.O., Reed, J.M., Kitchener, E.G., & Squire, L.R. (2003). Recognition memory and the human hippocampus. *Neuron*, 37(1), 171-180. doi:10.1016/S0896-6273(02)01147-9
- Maratos, E.J., Allan, K., & Rugg, M.D. (2000). Recognition memory for emotionally negative and neutral words: An ERP study. *Neuropsychologia*, 38(11), 1452-1465. doi:10.1016/S0028-3932(00)00061-0
- Marshall, S.P. (2002). The index of cognitive activity: measuring cognitive workload. In Proceedings from *The IEEE 7th Conference on Human Factors and Power Plants*, 7-5-7-9. doi:10.1109/HFPP.2002.1042860
- Marshall, S.P. (2007). Identifying cognitive state from eye metrics. *Aviation, Space, & Environmental Medicine*, 78(5,Suppl.), B165-B175.
- Martineau, J., Hernandez, N., Hiebel, L., Roché, L., Metzger, A., & Bonnet-Brilhault, F. (2011). Can pupil-size and pupil responses during visual scanning contribute to the diagnosis of autism spectrum disorder in children? *Journal of Psychiatric Research*, 45(8), 1077-1082. doi:10.1016/j.jpsychires.2011.01.008
- Maw, N.N., & Pomplun, M. (2004). Studying human face recognition with the gaze-contingent window technique. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the Twenty-Sixth Annual Meeting of the Cognitive Science Society*. Chicago, Illinois: Cognitive Science Society.
- McCarthy, G., & Wood, C.C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Clinical Neurophysiology*, 62(3), 203-208. doi:10.1016/0168-5597(85)90015-2
- McClelland, J.L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review*, 105(4), 724-760. doi:10.1037/0033-295X.105.4.734-760
- McElvain, J.L. (1970). *Pupillary dilations and galvanic skin response activity during competitive and noncompetitive paired association learning*. (Doctoral dissertation). Dissertation Abstracts International, 30, 3410B-3411B.
- McLaren, J.W., Erie, J.C., & Brubaker, R.F. (1992). Computerized analysis of pupillograms in studies of alertness. *Investigative Ophthalmology & Visual Science*, 33(3), 671-676.

- Mecklinger, A. (2000). Interfacing mind and brain: A neurocognitive model of recognition memory. *Psychophysiology*, 37(5), 565-582. doi:10.1111/1469-8986.3750565
- Mecklinger, A., von Cramon, D.Y., & Matthes-von Cramon, G. (1998). Event-related potential evidence for a specific recognition memory deficit in adult survivors of cerebral hypoxia. *Brain*, 121(10), 1919-1935. doi:10.1093/brain/121.10.1919
- Medina, J.J. (2008). The biology of recognition memory. *Psychiatric Times*. Retrieved March 14, 2009, from http://www.brainrules.net/pdf/JohnMedina_PsychTimes_June08.pdf
- Mentz, P. (1895). Die wirkung akustischer sinnesreize auf puls und athmung. *Philosophische Studien*, 11, 61-124.
- Merckelbach, H., & Christianson, S.A. (2007). Amnesia for homicides as a form of malingering. In S.A. Christianson (Ed.), *Offenders' memories of violent crimes* (pp. 165–190). Chichester, England: Wiley
- Merten, T., Thies, E., Schneider, K., & Stevens, A. (2009). Symptom validity testing in claimants with alleged posttraumatic stress disorder: Comparing the morel emotional numbing test, the structured inventory of malingered symptomatology, and the word memory test. *Psychological Injury & Law*, 2(3-4), 284-293. doi:10.1007/s12207-009-9057-0
- Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, 115(10), 2195-2222. doi:10.1016/j.clinph.2004.06.001
- Miller, E.K., & Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science*, 263(5146), 520-523. doi:10.1126/science.8290960
- Miller, G.A. (1956). The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97.
- Miller, L.J., Ryan, J.J., Carruthers, C.A., & Cluff, R.B. (2004). Brief screening indexes for malingering: A confirmation of Vocabulary minus Digit Span from the WAIS-III and the Rarely Missed Index from the WMS-III. *The Clinical Neuropsychologist*, 18, 327-333. doi:10.1080/13854040490501592
- Miller, N.R., & Newman, N.J. (Eds.) (2005). *Walsh & Hoyt's clinical neuro-ophthalmology: Volume one* (6th ed.). USA: Lippincott Williams & Wilkins.

- Miller, R.L. (1967). *The clinical validation of the pupillary response: The effect of chromatic and achromatic stimuli upon pupil responsivity*. (Unpublished doctoral dissertation). Michigan State University.
- Miller, S.R., & Donders, J. (2001). Subjective symptomology after traumatic head injury. *Brain Injury*, 15(4), 296-304. doi:10.1080/02699050118115
- Minassian, A., Granholm, E., Verney, S., & Perry, W. (2004). Pupillary dilation to simple versus complex tasks and its relationship to thought disturbance in schizophrenia patients. *International Journal of Psychophysiology*, 52(1), 53-62. doi:10.1016/j.ijpsycho.2003.12.008
- Mittenberg, W., Patton, C., Canyock, E.M., & Condit, D.C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical & Experimental Neuropsychology*, 24(8), 1094-1102. doi:10.1076/jcen.24.8.1094.8379
- Mohamed, F.B., Faro, S.H., Gordon, N.J., Platek, S.M., Ahmad, H., & Williams, J.M. (2006). Brain mapping of deception and truth telling about an ecologically valid situation: Functional MR imaging and polygraph investigation - Initial experience. *Neuroradiology*, 238(2), 679-688. doi:10.1148/radiol.2382050237
- Molnar, M. (1994). On the origin of the P3 event-related potential component. *International Journal of Psychophysiology*, 17(2), 129-144. doi:10.1016/0167-8760(94)90028-0
- Moloney, K.P., Jacko, J.A., Vidakovic, B., Sainfort, F., Leonard, V.K., & Shi, B. (2006). Leveraging data complexity: Pupillary behavior of older adults with visual impairment during HCI. *ACM Transactions on Computer-Human Interaction*, 13(3), 376-402. doi:10.1145/1183456.1183460
- Moray, N.E. (1979). *Mental workload: Its theory and measurement*. New York, Plenum Press.
- Mueller, D.J. (1970). Physiological techniques of attitude measurement. In G.F. Summers (Ed.), *Attitude measurement* (pp. 534-552). Chicago, IL: Rand McNally.
- Muldner, K., Christopherson, R., Atkinson, R., & Burleson, W. (2009). Investigating the utility of eye-tracking information on affect and reasoning for user modeling user modeling, adaptation, and personalization. In G.-J. Houben, G. McCalla, & F. Pianesi (Eds.), *Proceedings from The 17th International Conference on User Modeling, Adaptation, and Personalization* (pp. 138-149). Heidelberg, Germany: Springer-Verlag. doi:10.1007/978-3-642-02247-0_15

- Mulhall, E.F. (1915). Experimental studies in recall and recognition. *The American Journal of Psychology*, 26(2), 217-228.
- Müller-Jensen, A., & Hagenah, R. (1978). Simultaneous recording of pupillary hippus and EEG: Report of a case. *Journal of Neurology*, 217(3), 213-218.
doi:10.1007/BF00312964
- Murdock, B.B. (1982). A theory of the storage and retrieval of item and associative information. *Psychological Review*, 89(6), 609–626. doi:10.1037/0033-295X.89.6.609
- Murdock, B.B. (1985). An analysis of the strength-latency relationship. *Memory & Cognition*, 13(6), 511-521. doi:10.3758/BF03198322
- Murdock, B.B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review*, 104(4), 839-862.
- Murdock, B.B. (2006). Decision-making models of remember–know judgments: Comment on Rotello, Macmillan, and Reeder (2004). *Psychological Review*, 113(3), 648-656. doi:10.1037/0033-295x.113.3.648
- Murphy, P.R., Robertson, I.H., Balsters, J.H., & O'Connell, R.G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology*, 48(11), 1532-1543. doi:10.1111/j.1469-8986.2011.01226.x
- Myers, C., & Connor, M. (1992). Age differences in skill acquisition and transfer in an implicit learning paradigm. *Applied Cognitive Psychology*, 6(5), 429-442.
doi:10.1002/acp.2350060507
- Mylus, V., Braune, H.J., & Schepelmann, K. (2003). Dysfunction of the pupillary light reflex following migraine headache. *Clinical Autonomic Research*, 13(1), 16-21. doi:10.1007/s10286-003-0065-y
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral & Brain Sciences*, 13, 201-288.
- Naber, M., Rutishauser, U., & Einhäuser, W. (2011). Pupil size reflects familiarity and predicts later retrieval success during encoding of declarative memories. Unpublished manuscript. Philipps-Universität Marburg.
- Nakayama, M. (2006). Influence of blink on pupillary indices. In Proceedings from BioCAS: *The Biomedical Circuits and Systems Conference* (pp. 29-32). London: BioCAS. doi:10.1109/BIOCAS.2006.4600300

- Nakayama, M., & Shimizu, Y. (2004). Frequency analysis of task evoked pupillary response and eye-movement. In Proceedings from *The ACM Eye Tracking Research & Applications Symposium* (pp. 71-76). New York, NY: ACM Press. doi:10.1145/968363.968381
- Nakayama, M., Yasuike, I., & Shimizu, Y. (1990). Pupil-size changing by pattern brightness and pattern contents. *The Journal of the Institute of Television Engineers of Japan*, 44, 288-293.
- National Research Council. (2003). *The polygraph and lie detection*. Washington, DC: National Academies Press.
- Nave, C.R. (2010). Accommodation. In *Hyperphysics*. Retrieved from <http://hyperphysics.phy-astr.gsu.edu/hbase/vision/accom.html>
- Naveh-Benjamin, M., Craik, F.I.M., Guez, J., & Dori, H. (1998). The effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 24(5), 1091-1104.
- Neiss, R. (1988). Reconceptualizing arousal: Psychobiological states in motor performance. *Psychological Bulletin*, 103(3), 345-366. doi:10.1037/0033-2909.103.3.345
- Nessler, D., Mecklinger, A., & Penney, T.B. (2005). Perceptual fluency, semantic familiarity, and recognition-related familiarity: An electrophysiological exploration. *Cognitive Brain Research*, 22(2), 265-288. doi:10.1016/j.cogbrainres.2004.03.023
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J.D. (2005a). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*, 131(4), 510-532. doi:10.1037/0033-2909.131.4.510
- Nieuwenhuis, S., De Geus, E.J., & Aston-Jones, G. (2011a). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology*, 48(2), 162-175. doi:10.1111/j.1469-8986.2010.01057.x
- Nieuwenhuis, S., Forstmann, B.U., & Wagenmakers, E.-J. (2011b). Erroneous analyses of interactions in neuroscience: A problem of significance. *Nature Neuroscience*, 14(9), 1105-1107. doi:10.1038/nn.2886
- Nieuwenhuis, S., Gilzenrat, M.S., Holmes, B.D., & Cohen, J.D. (2005b). The role of the locus coeruleus in mediating the attentional blink: A neurocomputational

- theory. *Journal of Experimental Psychology: General*, 134(3), 291-307.
doi:10.1037/0096-3445.134.3.291
- Nissen, M.J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1), 1-32.
doi:10.1016/0010-0285(87)90002-8
- Nissen, M.J., Willingham, D., & Hartman, M. (1989). Explicit and implicit remembering: When is learning preserved in amnesia? *Neuropsychologia*, 27(3), 341-352. doi:10.1016/0028-3932(89)90023-7
- Norman, D.A., & Bobrow, D.J. (1975). On data limited and resource limited processes. *Cognitive Psychology*, 7(1), 44-64. doi:10.1016/0010-0285(75)90004-3
- Norman, E., Price, M., Duff, S., & Mentzoni, R. (2007). Gradations of awareness in a modified sequence learning task. *Consciousness and Cognition*, 16(4):809-837. doi:10.1016/j.concog.2007.02.004
- Norman, K.A., & O'Reilly, R.C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary learning systems approach. *Psychological Review*, 110(4), 611-646. doi:10.1037/0033-295X.110.4.611
- Norman, K.A., Tepe, K., Nyhus, E., & Curran, T. (2008). Event-related potential correlates of interference effects on recognition memory. *Psychonomic Bulletin & Review*, 15(1), 36-43. doi:10.3758/PBR.15.1.36
- Nunez, P.L. (1981). *Electric fields of the brain*. New York: Oxford University Press.
- Nunnally, C., Knott, P.D., Duchnowski, J., & Parker, K. (1967). Pupillary response as a general measure of activation. *Attention, Perception & Psychophysics*, 2(4), 149-155. doi:10.3758/BF03210310
- O'Bryant, S., & Lucas, J. (2006). Estimating the predictive value of the Test of Memory Malingering: An illustrative example for clinicians. *Clinical Neuropsychology*, 20(3), 533-540. doi:10.1080/13854040590967568
- O'Connor, M.K., & Ally, B.A. (2010). Using stimulus form change to understand memorial familiarity for pictures and words in patients with mild cognitive impairment and Alzheimer's disease. *Neuropsychologia*, 48(7), 2068-2074. doi:10.1016/j.neuropsychologia.2010.03.027
- Ohira, H. (1996). Eyeblick activity in a word-naming task as a function of semantic priming and cognitive load. *Perceptual & Motor Skills*, 82(3), 835-842. doi:10.2466/pms.1996.82.3.835

- Ohira, H., Winton, W.M., & Oyama, M. (1998). Effects of stimulus valence on recognition memory and endogenous eyeblinks: Further evidence for positive-negative asymmetry. *Personality & Social Psychology Bulletin*, 24(9), 986-993. doi:10.1177/0146167298249006
- Ohno, T., & Mukawa, N. (2004). A free-head, simple calibration, gaze tracking system that enables gaze-based interaction. In Proceedings from *The 2004 Symposium on Eye Tracking Research & Applications* (pp. 115-122). New York, NY: ACM. doi:10.1145/968363.968387
- Olichney, J.M., Van Petten, C., Paller, K.A., Salmon, D.P., Iragui, V.J., & Kutas, M. (2000). Word repetition in amnesia: Electrophysiological measures of impaired and spared memory. *Brain*, 123(9), 1948-1963. doi:10.1093/brain/123.9.1948
- Oliveira, F.T.P., Aula, A., & Russell, D.M. (2009). Discriminating the relevance of web search results with measures of pupil-size. In Proceedings from *The 27th International Conference on Human Factors in Computing Systems* (2209-2212). New York, NY: ACM Press. doi:10.1145/1518701.1519038
- O'Neill, W.D., Oroujeh, A.M., & Merritt, S.L. (1998). Pupil noise is a discriminator between narcoleptics and controls. *IEEE Transactions on Biomedical Engineering*, 45(3), 314-322. doi:10.1109/10.661156
- O'Neill, W.D., Oroujeh, A.M., Keegan, A.P., & Merritt, S.L. (1996). Neurological pupillary noise in narcolepsy. *Journal of Sleep Research*, 5(4), 265-271. doi:10.1111/j.1365-2869.1996.00265.x
- O'Neill, W.D., & Trick, K.P. (2001). The narcoleptic cognitive pupillary response. *IEEE Transactions on Biomedical Engineering*, 48(9), 963-968. doi:10.1109/10.942585
- Opitz, B. (2010). Context-dependent repetition effects on recognition memory. *Brain & Cognition*, 73(2), 110-118. doi:10.1016/j.bandc.2010.04.003
- Otero, S.C. (2010). *False memory in normal ageing*. (Unpublished doctoral dissertation). University of Sussex, England.
- Otero, S., Weekes, B., & Hutton, S. (2006). A novel association between pupil-size and recollective experience during recognition memory. In Proceedings from *The Second Biennial Conference on Cognitive Science*. Retrieved 12 May 2011 from http://www.cogsci.ru/cogsci06/index_e.htm

- Otero, S.C., Weekes, B.S., & Hutton, S.B. (2011). Pupil-size changes during recognition memory. *Psychophysiology*, 48(10), 1346-1353.
doi:10.1111/j.1469-8986.2011.01217.x
- Paivio, A. (1973). Psychophysiological correlates of imagery. In F.J. McGuigan, & R. Schoonover (Eds.), *The psychophysiology of thinking* (pp. 263-295). New York: Academic Press.
- Paivio, A., & Simpson, H.M. (1966). The effect of word abstractness and pleasantness on pupil-size during an imagery task. *Psychonomic Science*, 5(2), 55-56.
- Paivio, A., & Simpson, H.M. (1968). Magnitude and latency of the pupillary response during an imagery task as a function of stimuli abstractness and imagery ability. *Psychonomic Science*, 12(2), 45-46.
- Paller, K.A. (1990). Recall and stem-completion priming have different electrophysiological correlates and are modified differentially by directed forgetting. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 16(6), 1021-1032. doi:10.1037/0278-7393.16.6.1021
- Paller, K.A., Gonsalves, B., Grabowecky, M., Bozic, V.S., & Yamada, S. (2000). Electrophysiological correlates of recollecting faces of known and unknown individuals. *NeuroImage*, 11(2), 98-110. doi:10.1006/nimg.1999.0521
- Paller, K.A., & Kutas, M. (1992). Brain potentials during memory retrieval provide neurophysiological support of the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience*, 4, 375-391.
doi:10.1162/jocn.1992.4.4.375
- Paller, K.A., Kutas, M., & Mclsaac, H.K. (1995). Monitoring conscious recollection via the electrical activity of the brain. *Psychological Science*, 6(2), 107-111.
doi:10.1111/j.1467-9280.1995.tb00315.x
- Paller, K.A., Voss, J.L., & Boehm, S.G. (2007). Validating neural correlates of familiarity. *Trends in Cognitive Sciences*, 11(6), 243-250.
doi:10.1016/j.tics.2007.04.002
- Papesh, M.H., & Goldinger, S.D. (2011). Your effort is showing! Pupil dilation reveals memory heuristics. In P.A. Higham, & J.P. Leboe (Eds.), *Constructions of remembering and metacognition* (pp. 215-224). Hampshire: Palgrave Macmillan.

- Papesh, M.H., Goldinger, S.D., & Hout, M.C. (2011). Memory strength and specificity revealed by pupillometry. *International Journal of Psychophysiology*, 83(1), 56-64. doi:10.1016/j.ijpsycho.2011.10.002
- Park, H., Reder, L.M., & Dickison, D. (2005). The effects of word frequency and similarity on recognition judgments: The role of recollection. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 31(3), 568-578. doi:10.1037/0278-7393.31.3.568
- Parks, C.M., & Yonelinas, A.P. (2007). Moving beyond pure signal detection models: Comment on Wixted (2007). *Psychological Review*, 114(1), 188-202. doi:10.1037/0033-295X.114.1.188
- Partala, T., & Surakka, V. (2003). Pupil-size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, 59, 185-198. doi:10.1016/S1071-5819(03)00017-X
- Pavlov, I.P. (1927). *Conditioned reflexes*. Oxford: Clarendon Press.
- Payne, D.T., Parry, M.K., & Harasymiw, S.J. (1968). Percentage of pupillary dilation as a measure of item difficulty. *Perception & Psychophysics*, 4(3), 139-143. doi:10.3758/BF03210453
- Peavler, W.S. (1974). Pupil-size, information overload, and performance differences. *Psychophysiology*, 11(5), 559-566. doi:10.1111/j.1469-8986.1974.tb01114.x
- Peavler, W.S., & McLaughlin, J.P. (1967). The question of stimulus content and pupil-size. *Psychonomic Science*, 8(12), 505-506.
- Peter, G., Merten, T., Merckelbach, H., & Oswald, M. (2010). Detection of feigned crime-related amnesia: A multi-method approach. *Journal of Forensic Psychology Practice*, 10(5), 440-463. doi:10.1080/15228932.2010.489875
- Phillips, M.A., Bitsios, P., Szabadi, E., & Bradshaw, C.M. (2000a). Comparison of the antidepressants reboxetine, fluvoxamine and amitriptyline upon spontaneous pupillary fluctuations in healthy human volunteers. *Psychopharmacology*, 149(1), 72-76. doi:10.1007/s002139900334
- Phillips, M.A., Szabadi, E., & Bradshaw, C.M. (2000b). Comparison of the effects of clonidine and yohimbine on spontaneous pupillary fluctuations in healthy human volunteers. *Psychopharmacology*, 150(1), 85-89. doi: 10.1046/j.1365-2125.2000.00225.x
- Pivik, R.T., Broughton, R.J., Coppola, R., Davidson, R.J., Fox, N., & Nuwer, M.R. (1993). Guidelines for the recording and quantitative analysis of

- electroencephalographic activity in research contexts. *Psychophysiology*, 30(6), 547-588. doi:10.1111/j.1469-8986.1993.tb02081.x
- Pliny the Elder (1938). *Natural history with an English translation in ten volumes: volume III, libri VIII-XI* (H. Rackham, Trans.). London: William Heinemann. (Original work published 79 AD).
- Plonsey, R. (1963). Reciprocity applied to volume conductors and the EEG. *IEEE Transactions on Biomedical Engineering*, 19, 9-12.
- Polich, J., & Squire, L.E. (1993). P300 from amnesic patients with bilateral hippocampal lesions. *Electroencephalography & Clinical Neurophysiology*, 86(6), 408-417. doi:10.1016/0013-4694(93)90136-J
- Polt, J.M. (1970). Effect of threat of shock on pupillary response in a problem-solving task. *Perceptual & Motor Skills*, 31(2), 587-593. doi:10.2466/pms.1970.31.2.587
- Pomplun, M., & Sunkara, S. (2003). Pupil dilation as an indicator of cognitive workload in human-computer interaction. In D. Harris, V. Duffy, M. Smith, & C. Stephanidis (Eds.), *Proceedings from The 10th International Conference on Human-Computer Interaction - Human-centred computing: Cognitive, social, and ergonomic aspects, Vol. 3* (pp. 542-546). Crete, Greece: HCII.
- Pomplun, M., Sunkara, S., Fairley, A.V., & Xiao, M. (submitted). Using pupil-size as a measure of cognitive workload in video-based eye-tracking studies.
- Poock, G.K. (1973). Information processing vs. pupil diameter. *Perceptual & Motor Skills*, 37(3), 1000-1002.
- Porter, G., Tales, A., Trościąńko, T., Wilcock, G., Haworth, J., & Leonards, U. (2010). New insights into feature and conjunction search: I. Evidence from pupil-size, eye movements and ageing. *Cortex*, 46(5), 621-636. doi:10.1016/j.cortex.2009.04.013
- Porter, G., Trościąńko, T., & Gilchrist, I.D. (2007). Effort during visual search and counting: Insights from pupillometry. *The Quarterly Journal of Experimental Psychology*, 60(2), 211-229. doi:10.1080/17470210600673818
- Pothos, E.M. (2007). Theories of artificial grammar learning. *Psychological Bulletin*, 133(2), 227-244. doi:10.1037/0033-2909.133.2.227
- Pratt, R.W. (1970). Cognitive processing of uncertainty: Its effects on pupillary dilation and preference ratings. *Perception & Psychophysics*, 8(4), 193-198. doi:10.3758/BF03210204

- Pressman, M.R., Spielman, A.J., Korczyn, A.D., Rubenstein, A.E., Pollak, C.P., & Weitzman, E.D. (1984). Patterns of daytime sleepiness in narcoleptics and normals: A pupillometric study. *Electroencephalography & Clinical Neurophysiology*, 57(2), 129-133. doi:10.1016/0013-4694(84)90171-8
- Prime, S.L., Vesia, M., & Crawford, J.D. (2008). Transcranial magnetic stimulation over posterior parietal cortex disrupts trans-saccadic memory of multiple objects. *Journal of Neuroscience*, 28(27), 6938-6949. doi:10.1523/JNEUROSCI.0542-08.2008
- Pujol, M., & Kopelman, D. (2003). Psychogenic amnesia. *Practical Neurology*, 3(5), 292-299.
- Qiyuan, J., Richer, F., Wagoner, B.L., & Beatty, J. (1985). The pupil and stimulus probability. *Psychophysiology*, 22, 204-207. doi:10.1111/j.1469-8986.1985.tb01645.x
- Quian Quiroga, R., Mukamel, R., Isham, E.A., Malach, R., & Fried, I. (2008). Human single-neuron responses at the threshold of conscious recognition. *Proceedings of the National Academy of Sciences USA*, 105(9), 3599-3604. doi:10.1073/pnas.0707043105
- Quian Quiroga, R., & Kreiman, G. (2010). Postscript: About Grandmother cells and Jennifer Aniston neurons. *Psychological Reviews*, 117(1), 297-299. doi:10.1037/0033-295X.117.1.297
- Rabinowitz, J.C., & Graesser, A.C.I. (1976). Word recognition as a function of retrieval processes. *Bulletin of the Psychonomic Society*, 7, 75-77.
- Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition*, 21(1), 89-102.
- Rajaram, S., & Roediger, H.L., III. (1997). Remembering and knowing as states of consciousness during retrieval. In J.W. Schooler, & J.D. Cohen (Eds.), *Scientific approaches to consciousness. Carnegie Mellon symposia on cognition* (pp. 213-240). Mahwah, NJ: Erlbaum.
- Rajkowski, J., Kubiak, P., & Aston-Jones, G. (1993). Correlations between locus coeruleus neural activity, pupil diameter and behavior in monkey support a role of LC in attention. *Society for Neuroscience Abstracts*, 19, 974.
- Rajkowski, J., Kubiak, P., & Aston-Jones, G. (1994). Locus coeruleus activity in monkey: Phasic and tonic changes are associated with altered vigilance. *Brain Research Bulletin*, 35, 607-616. doi:10.1016/0361-9230(94)90175-9

- Rajkowski, J., Lu, W., Zhu, Y., Cohen, J.D., & Aston-Jones, G. (2000). Prominent projections from the anterior cingulate cortex to the locus coeruleus (LC) in rhesus monkey. *Society for Neuroscience Abstracts*, 26, 2230.
- Rajkowski, J., Majczynski, H., Clayton, E., & Aston-Jones, G. (2004). Activation of monkey locus coeruleus neurons varies with difficulty and performance in a target detection task. *Journal of Neurophysiology*, 92, 361-371. doi:10.1152/jn.00673.2003.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59-108. doi:10.1037/0033-295x.85.2.59
- Ratcliff, R., Clark, S.E., & Shiffrin, R.M. (1990). List-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 16(2), 163-178. doi:10.1037/0278-7393.16.2.163
- Ratcliff, R., & Murdock Jr., B.B. (1976). Retrieval processes in recognition memory. *Psychological Review*, 83(3), 190-214. doi:10.1037/0033-295X.83.3.190
- Ratcliff, R., Sheu, C.F., & Gronlund, S.D. (1992). Testing global memory models using ROC curves. *Psychological Review*, 99(3), 518-535. doi:10.1037/0033-295X.99.3.518
- Raymond, J.E., Shapiro, K.L., & Arnell, K.M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception & Performance*, 18(3), 849-860. doi:10.1037/0096-1523.18.3.849
- Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning & Verbal Behaviour*, 6(6), 855-863. doi:10.1016/S0022-5371(67)80149-X
- Reber, A.S. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology*, 81(1), 115-119. doi:10.1037/h0027454
- Reber, A.S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118(3), 219-235. doi:10.1093/acprof:oso/9780195106589.001.0001
- Reber, A.S., Walkenfeld, E., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 17(5), 888-896. doi:10.1037/0278-7393.17.5.888
- Recarte, M.A., & Nunes, L.M. (2003). Mental workload while driving: Effects on visual search, discrimination, and decision making. *Journal of Experimental Psychology/Applied*, 9(2), 119-137. doi:10.1037/1076-898X.9.2.119

- Recarte, M.A., Perez, E., Conchillo, A., & Nunes, L.M. (2008). Mental workload and visual impairment: Differences between pupil, blink, and subjective rating. *Spanish Journal of Psychology*, 11(2), 374-385.
- Reder, L.M., Nhouyvanisvong, A., Schunn, C.D., Ayers, M.S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 26(2), 294-320. doi:10.1037/0278-7393.26.2.294
- Reed, A.V. (1973). Speed-accuracy tradeoff in recognition memory. *Science*, 181(4099), 574-576. doi:10.1126/science.181.4099.574
- Reeves, A.G., & Swenson, R.S. (2004). Cranial nerves: Eye movements. In *Disorders of the nervous system: A primer* (Chapter 4). Dartmouth, NH: Dartmouth Medical School.
- Reynolds, G.D., & Richards, J.E. (2009). Cortical source localisation of infant cognition. *Developmental Neuropsychology*, 34(3), 312-329. doi:10.1080/87565640902801890
- Richer, F., & Beatty, J. (1985). Pupillary dilations in movement preparation and execution. *Psychophysiology*, 22(2), 204-207. doi:10.1111/j.1469-8986.1985.tb01587.x
- Richer, F., Silverman, C., & Beatty, J. (1983). Response selection and initiation in speeded reactions: A pupillometric analysis. *Journal of Experimental Psychology: Human Perception & Performance*, 9(3), 360-370, 1983. doi:10.1037/0096-1523.9.3.360.
- Roediger, H.L., & Byrne, J.H. (Eds.) (2008). *Learning and memory: A comprehensive reference* (Vol. 2). Oxford: Elsevier.
- Roediger, H.L.I., & McDermott, K.B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 21, 803-814.
- Roediger, H.L., Weldon, M.S., Stadler, M.L., & Riegler, G.L. (1992). Direct comparison of two implicit memory tests: Word fragment and word stem completion. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 18(6), 1251-1269. doi:10.1037/0278-7393.18.6.1251
- Rogers, R., & Bender, S.D. (2003) Evaluation of malingering and deception. In I.B. Weiner, & A.M. Goldstein (Eds.), *Handbook of psychology: Vol. 11. Forensic psychology* (pp. 109-129). New York: John Wiley & Sons, Inc.

- Rohrbaugh, J.W., Syndulko, K., & Lindsley, D.B. (1976). Brain wave components of the contingent negative variation in humans. *Science*, 191(4231), 1055-1057. doi:10.1126/science.1251217
- Rosenfeld, J.P., Ellwanger, J., & Sweet, J. (1995). Detecting simulated amnesia with event-related brain potentials. *International Journal of Psychophysiology*, 19(1), 1-11. doi:10.1016/0167-8760(94)00057-L
- Rosenfeld, J.P., Labkovsky, E., Winograd, M., Lui, M.A., Vandenboom, C., & Chedid, E. (2008). The Complex Trial Protocol (CTP): A new, countermeasure-resistant, accurate, P300-based method for detection of concealed information. *Psychophysiology*, 45, 906-919. doi:10.1111/j.1469-8986.2008.00708.x
- Rosenfeld, J.P., Soskins, M., Bosh, G., & Ryan, A., (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology*, 41(2), 205-219. doi:10.1111/j.1469-8986.2004.00158.x
- Ross, S.R., Millis, S.R., Krukowski, R.A., Putnam, S.H., & Adams, K.M. (2004). Detecting incomplete effort on the MMPI-2: An examination of the Fake Bad Scale in mild head injury. *Journal of Clinical & Experimental Neuropsychology*, 26(1), 115-124. doi:10.1076/jcen.26.1.115.23933
- Rotello, C.M., & Heit, E. (1999). Two-process models of recognition memory: Evidence for recall-to-reject? *Journal of Memory & Language*, 40(3), 432-453. doi:10.1006/jmla.1998.2623
- Rotello, C.M., & Heit, E. (2000). Associative recognition: A case of recall-to-reject processing. *Memory & Cognition*, 28(6), 907-922. doi:10.3758/BF03209339
- Rotello, C.M., & Macmillan, N.A. (2006). Remember-know models as decision strategies in two experimental paradigms. *Journal of Memory & Language*, 55, 479-494. doi:10.1016/j.jml.2006.08.002
- Rotello, C.M., Macmillan, N.A., & Reeder, J.A. (2004). Sum-difference theory of remembering and knowing: A two-dimensional signal-detection model. *Psychological Review*, 111(3), 588-616. doi:10.1037/0033-295X.111.3.588
- Rotello, C.M., Macmillan, N.A., & Van Tassel, G. (2000). Recall-to-reject in recognition: Evidence from ROC curves. *Journal of Memory & Language*, 43(1), 67-88. doi:10.1006/jmla.1999.2701
- Roubinovitch, J. (1900). Des variations du diamètre pupillaire en rapport avec l'effort intellectuel. In P. Janet (Ed.), *Quatrième congrès internationale de*

psychologie: Compte rendu des seances et texts des memoires, publies par les soins du Doctuer Pierre Janet. Paris: F Alcan.

- Rowe, D.W., Sibert, J., & Irwin, D. (1998). Heart rate variability: Indicator of user state as an aid to human-computer interaction. In Proceedings from *The SIGCHI Conference on Human Factors in Computing Systems*, 480-487. New York, NY: ACM Press. doi:10.1145/274644.274709
- Rowland, L.M., Thomas, M.L., Thorne, D.R., Sing, H.C., Krichmar, J.L., Davis, H.Q., ... Belenky, G. (2005). Oculomotor responses during partial and total sleep deprivation. *Aviation, Space, & Environmental Medicine*, 76(7S), C104-C113.
- Rugg, M.D. (1995). ERP studies of memory. In M.D. Rugg, & M.G.H. Coles (Eds.), *Electrophysiology of mind* (pp. 132-170). New York: Oxford University Press.
- Rugg, M.D., & Allan, K. (1999) Memory retrieval: An electrophysiological perspective. In M.S. Gazzaniga (Ed.) *The cognitive neurosciences* (pp.805-816). USA: MIT Press.
- Rugg, M.D., & Allan, K. (2000). Event-related potential studies of memory. In E. Tulving, & F.I.M. Craik (Eds.), *Oxford handbook of memory* (pp. 521-537). New York: Oxford University Press.
- Rugg, M.D., Cox, C.J.C., Doyle, M.C., & Wells, T. (1995). Event-related potentials and the recollection of low and high frequency words. *Neuropsychologia*, 33(4), 471-484. doi:10.1016/0028-3932(94)00132-9
- Rugg, M.D., & Doyle, M.C. (1992). Event-related potentials and recognition memory for low- and high-frequency words. *Journal of Cognitive Neuroscience*, 4(1), 69-79. doi:10.1162/jocn.1992.4.1.69
- Rugg, M.D., Mark, R.E., Walla, P., Schloerscheidt, A.M., Birch, C.S., & Allan, K. (1998a). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, 392(6676), 595-598. doi:10.1038/33396
- Rugg, M.D., & Nagy, M.E. (1989). Event-related potentials and recognition memory for words. *Electroencephalography & Clinical Neurophysiology*, 72(5), 395-406. doi:10.1016/0013-4694(89)90045-X
- Rugg, M.D., Pearl, S., Walker, P., Roberts, R.C., & Holdstock, J.S. (1994). Word repetition effects on event-related potentials in healthy young and old subjects, and in patients with Alzheimer-type dementia. *Neuropsychologia*, 32, 381-398. doi:10.1016/0028-3932(94)90085-X
- Rugg, M.D., Roberts, R.C., Potter, D.D., Pickles, C.D., & Nagy, M.E. (1991). Event-related potentials related to recognition memory. Effects of unilateral temporal

- lobectomy and temporal lobe epilepsy. *Brain*, 114(5), 2313-2332.
doi:10.1093/brain/114.5.2313
- Rugg, M.D., Schloerscheidt, A.M., & Mark, R.E. (1998b). An electrophysiological comparison of two indices of recollection. *Journal of Memory & Language*, 39(1), 47-69. doi:10.1006/jmla.1997.2555
- Rugg, M.D., & Yonelinas, A.P. (2003). Human recognition memory: A cognitive neuroscience perspective. *Trends in Cognitive Sciences*, 7(7), 313-319. doi:10.1016/S1364-6613(03)00131-1
- Rüsseler, J., Brett, A., Klaue, U., Sailer, M., & Münte, T.F. (2008). The effect of coaching on the simulated malingering of memory-impairment. *BioMed Central Neurology*, 8(37). doi:10.1186/1471-2377-8-37
- Rutishauser, U. (2008). *Learning and representation of declarative memories by single neurons in the human brain*. (Unpublished doctoral dissertation). California Institute of Technology.
- Rutishauser, U., Schuman, E.M., & Mamelak, A.N. (2008). Activity of human hippocampal and amygdala neurons during retrieval of declarative memories. *Proceedings for the National Academy of Sciences*, 105(1), 329-334. doi:10.1073/pnas.0706015105
- Sahraie, A., & Barbur, J.L. (1997). Pupil response triggered by the onset of coherent motion. *Graefes Archive for Clinical and Experimental Ophthalmology*, 235(8), 494-500. doi:10.1007/BF00947006.
- Samuel, R.Z., & Mittenberg, M.D. (2005). Determination of malingering in disability evaluations. *Primary Psychiatry*, 12(12), 60-68.
- Samuels, E.R., & Szabadi, E. (2008a). Functional neuroanatomy of the noradrenergic locus coeruleus: Its roles in the regulation of arousal and autonomic function part I: Principles of functional organisation. *Current Neuropsychopharmacology*, 6(3), 235-253. doi:10.1006/jmla.1997.2555
- Samuels, E.R., & Szabadi, E. (2008b). Functional neuroanatomy of the noradrenergic locus coeruleus: Its roles in the regulation of arousal and autonomic function part II: Physiological and pharmacological manipulations and pathological alterations of locus coeruleus activity in humans. *Current Neuropsychopharmacology*, 6(3), 254-285. doi:10.2174/157015908785777193
- Sanquist, T.F., Rohrbaugh, J.W., Syndulko, K., & Lindsley, D.B. (1980). Electrocortical signs of levels of processing: Perceptual analysis and

- recognition memory. *Psychophysiology*, 17(6), 568-576. doi:10.1111/j.1469-8986.1980.tb02299.x
- Sara, S.J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, 10(3), 211-223. doi:10.1038/nrn2573
- Schacter, D.L. (1977). EEG theta waves and psychological phenomena: A review and analysis. *Biological Psychology*, 5, 47-82. doi:10.1016/0301-0511(77)90028-X
- Schacter, D.L., & Buckner, R.L. (1998). On the relations among priming, conscious recollection, and intentional retrieval: evidence from neuroimaging research. *Neurobiology of Learning and Memory*, 70(1-2), 284-303. doi:10.1006/nlme.1998.3854
- Schacter, D.L., McAndrews, M.P., & Moscovitch, M. (1988). Access to consciousness: Dissociations between implicit and explicit knowledge in neuropsychological syndromes. In L. Weiskrantz (Ed.), *Thought without language* (pp. 242–278). Oxford: Oxford University Press.
- Schaefer Jr., T., Ferguson, J.B., Klein, J.A., & Rawson, E.B. (1968). Pupillary responses during mental activities. *Psychonomic Science*, 12(4), 137-138.
- Schagen, S., Schmand, B., de Sterke, S., & Lindeboom, J. (1997). Amsterdam Short-Term Memory Test: A new procedure for the detection of feigned memory deficits. *Journal of Clinical & Experimental Neuropsychology*, 19(1), 43-51. doi:10.1080/01688639708403835
- Schenk, K., & Sullivan, K.A. (2010). Do warnings deter rather than produce more sophisticated malingering? *Journal of Clinical & Experimental Neuropsychology*, 32(7), 752-762. doi:10.1080/13803390903512678
- Schiff, M. (1875). *La pupille considéré comme éthésiomètre*. Paris: J.B. Baillière.
- Schiff, M., & Foa, P. (1874). La pupille considérée comme esthésiomètre (R.G. de Choisy, Trans.). *Marseille Médical*, 2, 736-741.
- Schlemmer, K.B., Kulke, F., Kuchinke, L., & Van Der Meer, E. (2005). Absolute pitch and pupillary response: Effects of timbre and key color. *Psychophysiology*, 42(4), 465-472. doi:10.1111/j.1469-8986.2005.00306.x.
- Schloerscheidt, A.M., & Rugg, M.D., (2004). The impact of change in stimulus format on the electrophysiological indices of recognition. *Neuropsychologia*, 42, 451-466. doi:10.1016/j.neuropsychologia.2003.08.010
- Schluroff, M. (1982). Pupil responses to grammatical complexity of sentences. *Brain & Language*, 17(1), 133-145. doi:10.1016/0093-934X(82)90010-4

- Schluroff, M., Zimmermann, T.E., Freeman Jr., R.B., Hofmeister, K., Lorscheid, T., & Weber, A. (1986). Pupillary responses to syntactic ambiguity of sentences. *Brain & Language*, 27(2), 322-344. doi:10.1016/0093-934X(86)90023-4
- Schulman, A.I. (1967). Word length and rarity in recognition memory. *Psychonomic Science*, 9(4), 211-212.
- Schultheis, H., & Jameson, A. (2004). Assessing cognitive load in adaptive hypermedia systems: Physiological and behavioral methods. In P. De Bra, & W. Nejdl (Eds.), *Adaptive hypermedia and adaptive web-based systems* (pp. 225-234). Berlin: Springer Verlag.
- Schweitzer, N.M.J. (1956). Threshold measurements of the pupil the dark adapted eye. *Documenta Ophthalmologica*, 10(1), 1-78. doi:10.1007/BF00172098
- Scott, R.B., & Dienes, Z. (2008). The conscious, the unconscious, and familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 34(5), 1264-1288. doi:10.1037/a0012943
- Scott, R.B., & Dienes, Z. (2009). Knowledge applied to new domains: The unconscious succeeds where the conscious fails. *Consciousness & Cognition*, 19(1), 391-398. doi:10.1016/j.concog.2009.11.009
- Scott, T.R., Wells, W.H., Wood, D.Z., & Morgan, D.I. (1967). Pupillary response and sexual interest reexamined. *Journal of Clinical Psychology*, 23(4), 433-438. doi:10.1002/1097-4679(196710)23:4<433::AID-JCLP2270230408>3.0.CO;2-2
- Shanks, D.R., & St.John, M.F. (1994). Characteristics of dissociable human learning systems. *Behavioral & Brain Sciences*, 17(3), 367-448. doi:10.1017/S0140525X00035032
- Shepard, R.N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning & Verbal Behavior*, 6(1), 156-163. doi:10.1016/S0022-5371(67)80067-7
- Shiffrin, R.M., & Steyvers, M. (1997). A model of recognition memory: REM – Retrieving effectively from memory. *Psychological Bulletin & Review*, 4(2), 145-166.
- Shiffrin, R.M., & Steyvers, M. (1998). The effectiveness of retrieval from memory. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 73-95). London: Oxford University Press.
- Shore, D.I., & Klein, R.M. (2000). On the manifestations of memory in visual search. *Spatial Vision*, 14(1), 59-75. doi:10.1163/156856801741369

- Siegle, G.J., Granholm, E., Ingram, R.E., & Matt, G.E. (2001). Pupillary and reaction time measures of sustained processing of negative information in depression. *Biological Psychiatry*, 49(7), 624-636.
- Siegle, G.J., Ichikawa, N., & Steinhauer, S. (2008). Blink before and after you think: blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology*, 45(5), 679-687. doi:10.1111/j.1469-8986.2008.00681.x
- Siegle, G.J., Steinhauer, S.R., Carter, C.S., Ramel, W., & Thase, M.E. (2003a). Do the seconds turn into hours? Relationships between sustained pupil dilation in response to emotional information and self-reported rumination. *Cognitive Therapy & Research*, 27(3), 365-382. doi:10.1023/A:102397 4602357
- Siegle, G.J., Steinhauer, S.R., Carter, C.S., & Thase, M.E. (submitted). Is sustained processing specific to emotional information in depression? Evidence from pupil dilation.
- Siegle, G.J., Steinhauer, S.R., Stenger, V.A., Konecky, R., & Carter, C.S. (2003b). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *Neuroimage*, 20, 114-124. doi:10.1016/S1053-8119(03)00298-2
- Siegle, G.J., Steinhauer, S.R., & Thase, M.E. (2004). Pupillary assessment and computational modeling of the stroop task in depression. *International Journal Psychophysiology*, 52, 63-76. doi:10.1016/j.ijpsycho.2003.12.010
- Silk, J.S., Dahl, R.E., Ryan, N.D., Forbes, E.E., Axelson, D.A., Birmaher, B., & Siegle, G.J. (2007). Pupillary reactivity to emotional information in child and adolescent depression: Links to clinical and ecological measures. *American Journal of Psychiatry*, 164(12), 1873–1880. doi:10.1176/appi.ajp.2007.06111816
- Silk, J.S., Siegle, G.J., Whalen, D.J., Ostapenko, L.J., Ladouceur, C.D., & Dahl, R.E. (2009). Pubertal changes in emotional information processing: pupillary, behavioral, and subjective evidence during emotional word identification. *Developmental Psychopathology*, 21(1), 7-26. doi:10.1017/S0954579409000029
- Simpson, H.M. (1969). Effects of a task-relevant response on pupil-size. *Psychophysiology*, 6(2), 115-121. doi:10.1111/j.1469-8986.1969.tb02890.x

- Simpson, H.M., & Climan, M.H. (1971). Pupillary and electromyographic changes during an imagery task. *Psychophysiology*, 8(4), 483-490. doi:10.1111/j.1469-8986.1971.tb00480.x
- Simpson, H.M., & Hale, S.M. (1969). Pupillary changes during a decision-making task. *Perceptual & Motor Skills*, 29(2), 495-498. doi:10.2466/pms.1969.29.2.495
- Simpson, H.M., & Molloy, F.M. (1971). Effects of audience anxiety on pupil-size. *Psychophysiology*, 8, 491-496. doi:10.1111/j.1469-8986.1971.tb00481.x
- Simpson, H.M., Molloy, F.M., Hale, S.M., & Climan, M.H. (1968). Latency and magnitude of the pupillary response during an imagery task. *Psychonomic Science*, 13, 293-294.
- Simpson, H.M., & Paivio, A. (1966). Changes in pupil-size during an imagery task without motor response involvement. *Psychonomic Science*, 5, 405-406.
- Simpson, H.M., & Paivio, A. (1968). Effects on pupil-size of manual and verbal indicators of cognitive task fulfillment. *Perception & Psychophysics*, 3(3A), 185-190. doi:10.3758/BF03212726
- Skoglund, C.R. (1943). On the influence of alcohol on the pupillary light reflex in man. *Acta Physiologica Scandinavica*, 6(1), 94-96. doi:10.1111/j.1748-1716.1943.tb02831.x
- Slick, D., Hopp, G., Strauss, E., & Thompson, G.B. (1997). *VSVT: Victoria Symptom Validity Test (Version 1.0)*. Odessa, Florida: Psychological Assessment Resources.
- Slick, D.J., Sherman, E.M.S., & Iverson, G.L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, 13(4), 545-561. doi:10.1076/1385-4046(199911)13:04;1-Y;FT545
- Slooter, J., & van Norren, D. (1980). Visual acuity measured with pupil responses to checkerboard stimuli. *Investigative Ophthalmology & Visual Science*, 19(1), 105-108.
- Smith, G.P., & Burger, G.K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology. *Journal of the American Academy of Psychiatry & the Law*, 25(2), 183-189.
- Smith, M.E. (1993). Neurophysiological manifestations of recollective experience during recognition memory judgments. *Journal of Cognitive Neuroscience*, 5(1), 1-13. doi:10.1162/jocn.1993.5.1.1

- Smith, M.E., & Guster, K. (1993). Decomposition of recognition memory event-related potentials yields target, repetition, and retrieval effects. *Electroencephalography & Clinical Neurophysiology*, 86(5), 335-343. doi:10.1016/0013-4694(93)90046-X
- Smith, M.E., & Halgren, E. (1989). Dissociation of recognition memory components following temporal lobe lesions. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 15(1), 50-60. doi:10.1037/0278-7393.15.1.50
- Smith, M.E., Halgren, E., Sokolik, M., Baudena, P., Musolino, A., Liegois-Chauvel, C., & Chauvel, P. (1990). The intracranial topography of the P3 event-related potential elicited during auditory oddball. *Electroencephalography & Clinical Neurophysiology*, 76(3), 235-248. doi:10.1016/0013-4694(90)90018-F
- Smith, M.E., Stapleton, J.M., & Halgren, E. (1986). Human medial temporal lobe potentials evoked in memory and language tasks. *Electroencephalography & Clinical Neurophysiology*, 63(2), 145-159. doi:10.1016/0013-4694(86)90008-8
- Soderlund, H., Parker, E.S., Schwartz, B.L., & Tulving, E. (2005). Memory encoding and retrieval on the ascending and descending limbs of the blood alcohol concentration curve. *Psychopharmacology*, 182(2), 305-317. doi:10.1007/s00213-005-0096-2
- Sokolov, E.N. (1963). *Perception and the conditioned reflex*. Oxford: Pergamon Press.
- Sokolski, K.N., Nguyen, B.D., & DeMet, E.M. (2000). Decreases in dilated pupil-size in depressed patients with age may reflect adrenergic changes. *Psychiatry Research*, 94(3), 267-272.
- Spector, R.H. (1990). The pupils. In H.K. Walker, W.D. Hall, & J.W. Hurst (Eds.), *Clinical methods: The history, physical, and laboratory examinations* (3rd ed., pp. 300-304). Boston: Butterworths.
- Spence, S.A., Farrow, T.F.D., Herford, A.E., Wilkinson, I.D., Zheng, Y., & Woodruff, P.W.R. (2001). Behavioural and functional anatomical correlates of deception in humans. *NeuroReport*, 12(13), 2849-2853.
- Spencer, K.M., Vila Abad, E., & Donchin, E. (2000). On the search for the neurophysiological manifestation of recollective experience. *Psychophysiology*, 37(4), 494-506. doi:10.1111/1469-8986.3740494
- Squire, L.R., Wixted, J.T., & Clark, R.E. (2007). Recognition memory and the medial temporal lobe: A new perspective. *Nature*, 445(7130), 905-912. doi:10.1038/nrn2154

- Srinivasan, R., Tucker, D.M., & Murias, M. (1998). Estimating the spatial nyquist of the human EEG. *Behavioral Research Methods*, 30(1), 8-19.
doi:10.3758/BF03209412
- Staal, M.A. (2004). *Stress, cognition, and human performance: A literature review and conceptual framework* (Tech. Memorandum 212824). Moffett Field, CA: National Aeronautics and Space Administration Ames Research Center.
- Stanners, R.F., Headley, D.B., & Clark, W.R. (1972). The pupillary response to sentences: Influences of listening set and deep structure. *Journal of Verbal Learning & Verbal Behavior*, 11(2), 257-263. doi:10.1016/S0022-5371(72)80086-0
- Stark, C.E., & Squire, L.R. (2003). Hippocampal damage equally impairs memory for single items and memory for conjunctions. *Hippocampus*, 13(2), 281-292.
doi:10.1002/hipo.10085
- Stark, C.E.L., Bayley, P.J., & Squire, L.R. (2002). Recognition memory for single items and for associations is similarly impaired following damage to the hippocampal region. *Learning & Memory*, 9(5), 238-242.
doi:10.1101/lm.51802
- Steinhauer, S.R. (1982). *Emitted and evoked pupillary responses and event-related potentials as a function of reward and task involvement*. (Unpublished doctoral dissertation). City University of New York, NY, USA.
- Steinhauer, S.R. (2002). Pupillary responses, cognitive psychophysiology and psychopathology. Retrieved 10 June 2007 from
<http://www.wpic.pitt.edu/research/biometrics/Publications/Biometrics%20Archives%20PDF/874-2002%20Steinhauer%20PUPILWEB.pdf>
- Steinhauer, S.R., Condray, R., & Kasperek, A. (2000). Cognitive modulation of midbrain function: Task-induced reduction of the pupillary light reflex. *International Journal of Psychophysiology*, 39(1), 21-30. doi:10.1016/S0167-8760(00)00119-7
- Steinhauer, S.R., & Hakerem, G. (1992). The pupillary response in cognitive psychophysiology and schizophrenia. In D. Friedman, & G. Bruder (Eds.), *Psychophysiology and experimental psychopathology: A tribute to Samuel Sutton* (pp. 182-204). New York, NY: Academy of Sciences.
- Steinhauer, S.R., Hakerem, G., & Spring, B.J. (1979). The pupillary response as a potential indicator of vulnerability to schizophrenia. *Psychopharmacology Bulletin*, 15(1), 44-45.

- Steinhauer, S.R., Siegle, G.J., Condray, R., & Pless, M. (2004). Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *International Journal of Psychophysiology*, 52(1), 77-86.
doi:10.1016/j.ijpsycho.2003.12.005
- Stelmack, R.M., & Leckett, W.J. (1974). Effect of artificial pupil-size on recognition threshold. *Perceptual & Motor Skills*, 39(2), 739-742.
doi:10.2466/pms.1974.39.2.739
- Stelmack, R.M., & Mandelzys, N. (1975). Extraversion and pupillary response to affective and taboo words. *Psychophysiology*, 12(5), 536-540.
doi:10.1111/j.1469-8986.1975.tb00042.x
- Steptoe, W., Steed, A., Rovira, A., & Rae, J. (2010). Lie tracking: Social presence, truth and deception in avatar-mediated telecommunication. In Proceedings from *The Conference on Human Factors in Computing Systems*, 2, 1039-1048. Atlanta, Georgia: ACM. doi:10.1145/1753326.1753481
- Stern, J.A., Walrath, L.C., & Goldstein, R. (1984). The endogenous eyeblink. *Psychophysiology*, 21(1), 22-33. doi:10.1111/j.1469-8986.1984.tb02312.x
- Stern, P.C. (2003). The polygraph and lie detection. In *Report of the National Research Council Committee to Review the Scientific Evidence on the Polygraph* (pp. 11-28). Washington, DC: The National Academies Press.
- Sternback, R.A. (1966). *Principles of psychophysiology*. New York: Academic Press.
- Stevens, W.D., Wig, G.S., & Schacter, D.L. (2008). Implicit memory and priming. In H.L. Roediger III, & J.H. Byrne (Eds.), *Learning and memory, volume 2: Cognitive psychology of memory* (pp. 623-644). London: Elsevier.
- Stone, G.M., & Rothenheber, E. (1992). Advanced scientific detection of deception - ERP augmented polygraphy. In Proceedings from *The International Carnahan Conference on Security Technology: Crime Countermeasures* (pp 72-73). Atlanta, GA: Institute of Electrical and Electronics Engineers.
doi:10.1109/CCST.1992.253758
- Sullivan, K.N., Manfredi, F., & Behnke, R.H. (1968). Hippus in Cheyne-Stokes respiration: Observations in three patients with rhythmic respiratory and pupillary changes. *Archives of Internal Medicine*, 122(2), 116-121.
- Suzuki, M., Johnson, J.D., & Rugg, M.D. (2011). Decrements in hippocampal activity with item repetition during continuous recognition: An fMRI study. *Journal of Cognitive Neuroscience*, 23(6), 1522-1532. doi:10.1162/jocn.2010.21535

- Swihart, A.A, Harris, K.M., & Hatcher, L.L. (2008). Inability of the Rarely Missed Index to identify simulated malingering under more realistic assessment conditions. *Journal of Clinical & Experimental Neuropsychology*, 30(1), 120-126. doi:10.1080/13803390701249044
- Szabadi, E., & Bradshaw, C.M. (1996). Autonomic pharmacology of $\alpha 2$ -adrenoceptors. *Journal of Psychopharmacology*, 10(3), 6-18.
- Tardif, H.P., Barry, R.J., Fox, A.M., & Johnstone, S.J. (2000). Detection of feigned recognition memory-impairment using the old/new effect of the event-related potential. *International Journal of Psychophysiology*, 36(1), 1-9. doi:10.1016/S0167-8760(00)00083-0
- Thackray, R.I. (1969) *Patterns of physiological activity accompanying performance in a perceptual motor task* (Federal Aviation Administration Report. No. 69-8). Washington DC, USA: Office of Aviation Medicine.
- Tipper, S.P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology*, 37(4), 571-590. doi:10.1080/14640748508400920
- Tombaugh, T.N. (1996). *Test of Memory and Malingering*. North Tonawanda, NY: Multi-Health Systems.
- Trott, C.T., Friedman, D., Ritter, W., & Fabiani, M. (1997). Item and source memory: Differential age effects revealed by event-related potentials. *NeuroReport*, 8(15), 3373-3378.
- Trott, C.T., Friedman, D., Ritter, W., Fabiani, M., & Snodgrass, J.G. (1999). Episodic priming and memory for temporal source: Event-related potentials reveal age-related differences in prefrontal functioning. *Psychology & Aging*, 14(3), 390-413. doi:10.1037/0882-7974.14.3.390
- Tryon, W.W. (1975). Pupillometry: A survey of sources of variation. *Psychophysiology*, 12(1), 90-93. doi:10.1111/j.1469-8986.1975.tb03068.x
- Tsivilis, D., Otten, L.J., & Rugg, M.D. (2001). Context effects on the neural correlates of recognition memory: An electrophysiological study. *Neuron*, 31(3), 497-505. doi:10.1016/S0896-6273(01)00376-2
- Tucker, D.M. (1993). Spatial sampling of head electrical fields: The geodesic sensor net. *Electroencephalography & Clinical Neurophysiology*, 87(3), 154-163. doi:10.1016/0013-4694(93)90121-B

- Tucker, D.M., Liotti, M., Potts, G.F., Russell, G.S., & Posner, M.I. (1994). Spatiotemporal analysis of brain electrical fields. *Human Brain Mapping*, 1(2), 134-152. doi:10.1002/hbm.460010206
- Tugcu, H. (2010). Forensic medicine and malingering. *Turkiye Klinikleri Journal of Medical Sciences*, 30(4S), S86-S86.
- Tulving, E. (1983). *Elements of episodic memory*. New York, USA: Oxford University Press.
- Tulving, E. (1985a). Memory and consciousness. *Canadian Psychologist*, 26, 1-12.
- Tulving, E. (1985b). How many memory systems are there? *American Psychologist*, 40(4), 385-398. doi:10.1037/0003-066X.40.4.385
- Tulving, E., & Kroll, N. (1995). Novelty assessment in the brain and long-term memory encoding. *Psychonomic Bulletin & Review*, 2(3), 387-390. doi:10.3758/BF03210977
- Tulving, E., & Schacter, D.L. (1990). Priming and human memory systems. *Science*, 247(4940), 301-306. doi:10.1126/science.2296719
- Turriziani, P., Oliveri, M., Salerno, S., Costanzo, F., Koch, G., Caltagirone, C., & Carlesimo, G.A. (2008). Recognition memory and prefrontal cortex: Dissociating recollection and familiarity processes using rTMS. *Behavioural Neurology*, 19(1-2), 23-27.
- Ucar, M., & Atac, A. (2010). The term of malingering in terms of medical ethics: The ethical dimension of patient's attitudes and behaviours intended for misleading physician. *Turkiye Klinikleri Journal of Medical Sciences*, 30(4S), S1-S6.
- Uhl, F., Franzen, P., Serles, W., Lang, W., Lindinger, G., & Deecke, L. (1990). Anterior frontal cortex and the effect of proactive interference in paired associate learning. A DC-potential study. *Journal of Cognitive Neuroscience*, 2(4), 373-382. doi:10.1162/jocn.1990.2.4.373
- Ukai, K. (1985). Spatial pattern as a stimulus to the pupillary system. *Journal of the Optical Society of America*, 2(7), 1094-1099. doi:10.1364/JOSAA.2.001094
- Urbach, T.P., & Kutas, M. (2002). The intractability of scaling scalp distribution to infer neuroelectric sources. *Psychophysiology*, 39(6), 791-808. doi:10.1017/S0048577202010648
- Urry, H.L., van Reekum, C.M., Johnstone, T., Kalin, N.H., Thurow, M.E., Schaefer, H.S., . . . Davidson, R.J. (2006). Amygdala and ventromedial prefrontal cortex are inversely coupled during regulation of negative affect and predict the

- diurnal pattern of cortisol secretion among older adults. *Journal of Neuroscience*, 26(16), 4415-4425. doi:10.1523/JNEUROSCI.3215-05.2006
- Usher, M., Cohen, J.D., Servan-Schreiber, D., Rajkowski, J., & Aston-Jones, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science*, 283(5401), 549-554. doi:10.1126/science.283.5401.549
- Vagnini, V.L., Berry, D.T.R., Clark, J.A., & Jiang, Y. (2008). New measures to detect malingered neurocognitive deficit: Applying reaction time and event-related potentials. *Journal of Clinical & Experimental Neuropsychology*, 30(7), 766-776. doi:10.1080/13803390701754746
- Van Droof, L.V., Ledoux, K., Pickett, E.J., Buz, E., Billings, N., & Gordon, B. (2010). Advancing the measurement of receptive language in nonverbal individuals with autism. Presented at *The 2010 International Meeting for Autism Research, Philadelphia, PA*.
- van Hooff, J.C., Brunia, C.H.M., & Allen, J.J.B. (1996). Event-related potentials as indirect measures memory. *International Journal of Psychophysiology*, 21(1), 15-31. doi:10.1016/0167-8760(95)00043-7
- van Hooff, J.C., Sargeant, E., Foster, J.K., & Schmand, B.A. (2009). Identifying deliberate attempts to fake memory-impairment through the combined use of reaction time and event-related potential measures. *International Journal of Psychophysiology*, 73(3), 246-256. doi:10.1016/j.ijpsycho.2009.04.002
- Van Orden, K.F., Limbert, W., Makeig, S., & Jung, T.P. (2001). Eye activity correlates of workload during a visuospatial memory task. *Human Factors: The Journal of the Human Factors & Ergonomics Society*, 43(1), 111-121. doi:10.1518/001872001775992570
- Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, 18(4), 380-393. doi:10.3758/BF03197127
- Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory & Cognition*, 19(1), 95-112. doi:10.3758/BF03198500
- Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & McIsaac, H. (1991). Fractionating the word repetition effect with event related potentials. *Journal of Cognitive Neuroscience*, 3(2), 131-150. doi:10.1162/jocn.1991.3.2.131
- Van Rijn, H., Dalenberg, J., Borst, J., & Sprenger, S.A. (submitted). Memory strength correlates negatively with pupil dilation in a delayed paired-associate task.

- Van Turennout, M.V., Ellmore, T., & Martin, A. (2000). Long-lasting cortical plasticity in the object naming system. *Nature Neuroscience*, 3(12), 1329-1334. doi:10.1038/81873
- Verfaellie, M., Bauer, R.M., & Bowers, D. (1991). Autonomic and behavioral evidence of "implicit" memory in amnesia. *Brain & Cognition*, 15(1), 10-25. doi:10.1016/0278-2626(91)90012-W
- Verleger, R. (1995). Memory-related EEG potentials: Slow negativities, priming positivity, recognition positivity, and Dm: Commentary on Klimesch on Memory-Brain. *Psychology*, 6(27), 3.
- Verney, S.P., Granholm, E., Marshall, S.P., Malcarne, V.L., & Saccuzzo, D.P. (2005). Culture-fair cognitive ability assessment: Information processing and psychophysiological approaches. *Assessment*, 12(3), 303-319. doi:10.1177/1073191105276674
- Vilberg, K.L., & Rugg, M.D. (2008). Memory retrieval and the parietal cortex: A review of evidence from a dual-process perspective. *Neuropsychologia*, 46(7), 1787-1799. doi:10.1016/j.neuropsychologia.2008.01.004
- Võ, M.L.H., Jacobs, A.M., Kuchinke, L., Hofmann, M., Conrad, M., Schacht, A., & Hutzler, F. (2008). The coupling of emotion and cognition in the eye: Introducing the pupil old/new effect. *Psychophysiology*, 45(1), 130-140. doi:10.1111/j.1469-8986.2007.00606.x
- Voss, J.L., & Paller, K.A. (2008). Brain substrates of implicit and explicit memory: The importance of concurrently acquired neural signals of both memory types. *Neuropsychologia*, 46, 3021-3029. doi:10.1016/j.neuropsychologia.2008.07.010
- Wagner, A.D., Desmond, J.E., Demb, J.B., Glover, G.H., & Gabrieli, J.D.E. (1997). Semantic repetition priming for verbal and pictorial knowledge: A functional MRI study of left inferior prefrontal cortex. *Journal of Cognitive Neuroscience*, 9(6), 714-726. doi:10.1162/jocn.1997.9.6.714
- Wagner, A.D., Maril, A., & Schacter, D.L. (2000). Interactions between forms of memory: When priming hinders new learning. *Journal of Cognitive Neuroscience*, 12(S2), 52-60. doi:10.1162/089892900564064
- Wais, P.E., Wixted, J.T., Hopkins, R.O., & Squire, L.R. (2006). The hippocampus supports both the recollection and the familiarity components of recognition memory. *Neuron*, 49(3), 459-466. doi:10.1016/j.neuron.2005.12.020

- Walzl, M., Hagen, R., & Prummer, K. (2007). Pupillometric examinations for sleepiness on professional drivers. *Pupillometrische untersuchungen auf schläfrigkeit bei berufskraftfahrern*, 57(11), 349-364.
- Wang, J.T. (2010). Pupil dilation and eye tracking. In M. Schulte-Mecklenbeck, A. Kuehberger, & R. Ranyard (Eds.), *Handbook of process-tracing methods* (pp. 185-204). Hove: Psychology Press.
- Wang, J.T., Spezio, M., & Camerer, C.F. (2010). Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *American Economic Review*, 100(3), 984-1007. doi:10.1257/aer.100.3.984
- Warren, L.R. (1980). Evoked-potential correlates of recognition memory. *Biological Psychology*, 11(1), 21-35. doi:10.1016/0301-0511(80)90023-X
- Wasserman, S., & Bockenholt, U. (1989). Bootstrapping: Applications to psychophysiology. *Psychophysiology*, 26(2), 208-221. doi:10.1111/j.1469-8986.1989.tb03159.x
- Webb, A.K., Honts, C.R., Kircher, J.C., Bernhardt, P., & Cook, A.E. (2009). Effectiveness of pupil diameter in a probable-lie comparison question test for deception. *Legal & Criminological Psychology*, 14, 279-292. doi:10.1348/135532508X398602
- Wechsler, D. (1997). *Wechsler Memory Scale – third edition*. San Antonio, TX: The Psychological Corporation.
- Weiermann, B., Stephan, M.A., Kaelin-Lang, A., & Meier, B. (2010). Is there a recognition memory deficit in Parkinson's Disease? Evidence from estimates of recollection and familiarity. *International Journal of Neuroscience*, 120(3), 211-216. doi:10.3109/00207450903506510
- Wenger, W., & Videbeck, R. (1969). Eye pupillary measurement of aesthetic responses to forest scenes. *Journal of Leisure Research*, 1(2), 149-161.
- Westerman, D.L. (2001). The role of familiarity in item recognition, associative recognition, and plurality recognition on self-paced and speeded tests. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 27(3), 723-732. doi:10.1037/0278-7393.27.3.723
- Wheeler, M.E., & Buckner, R.L. (2004). Functional-anatomic correlates of remembering and knowing. *NeuroImage*, 21(4), 1337-1349. doi:10.1016/j.neuroimage.2003.11.001

- White, G.L., & Maltzman, I. (1978). Pupillary activity while listening to verbal passages. *Journal of Research in Personality*, 12(3), 361-369.
doi:10.1016/0092-6566(78)90062-4
- Wickelgren, W.A. (1979). Chunking and consolidation: A theoretical synthesis of semantic networks, configuring in conditioning, S-R versus cognitive learning, normal forgetting, the amnesic syndrome, and the hippocampal arousal system. *Psychological Review*, 86(1), 44-60. doi:10.1037/0033-295X.86.1.44
- Wiese, H., & Daum, I. (2006). Frontal positivity discriminates true from false recognition. *Brain Research*, 1075, 183-192. doi:10.1016/j.brainres.2005.12.117
- Wiggs, C.L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, 8(2), 227-233. doi:10.1016/S0959-4388(98)80144-X
- Wilder, J. (1958). Modern psychophysiology and the law of initial value. *American Journal of Psychotherapy*, 12(2), 199-221.
- Wilding, E.L., Doyle, M.C., & Rugg, M.D. (1995). Recognition memory with and without retrieval of context: an event-related potential study. *Neuropsychologia*, 33(6), 743-767. doi:10.1016/0028-3932(95)00017-W
- Wilding, E.L., & Rugg, M.D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Brain*, 119(3), 889-905.
doi:10.1093/brain/119.3.889
- Wilding, E.L., & Rugg, M.D. (1997a). Event-related potential and the recognition memory exclusion task. *Neuropsychologia*, 35(2), 119-128.
doi:10.1016/S0028-3932(96)00076-0
- Wilding, E.L., & Rugg, M.D. (1997b). An event-related potential study of recognition memory for words spoken aloud or heard. *Neuropsychologia*, 35(9), 1185-1195. doi:10.1016/S0028-3932(97)00048-1
- Wilding, E.L., & Sharpe, H. (2003). Episodic memory encoding and retrieval: Recent insights from event-related potentials. In A. Zani, & A.M. Proverbio (Eds.), *The cognitive electrophysiology of mind and brain* (pp. 169-196). San Diego, CA: Academic Press.
- Wilhelm, B.J., Heine, C., Widmann, A., Durst, W., Lüdtke, H., & Otto, G. (2010). How sleepy are construction workers during extended shifts? An objective investigation of spontaneous pupillary oscillations in tunnel workmen. *Somnologie*, 14(3), 200-206. doi:10.1007/s11818-010-0480-y

- Wilhelm, B.J., Widmann, A., Durst, W., Heine, C., & Otto, G. (2009). Objective and quantitative analysis of daytime sleepiness in physicians after night duties. *International Journal of Psychophysiology*, 72(3), 307-313. doi:10.1016/j.ijpsycho.2009.01.008
- Wilks, S. (1883). On the pupil in emotional states. *Brain*, 6(1), 1-6. doi:10.1093/brain/6.1.1
- Wilks, S. (1885). The pupil of the eyes during emotion. *Nature*, 31(803), 458. doi:10.1038/031458c0
- Winograd, M.R., & Rosenfeld, J.P. (2011). Mock crime application of the Complex Trial Protocol (CTP) P300-based concealed information test. *Psychophysiology*, 48, 155-161. doi:10.1111/j.1469-8986.2010.01054.x
- Wixted, J.T. (2007a). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, 114(1), 152-176. doi:10.1037/0033-295X.114.1.152
- Wixted, J.T. (2007b). Spotlighting the probative findings: A reply to Parks and Yonelinas (2007). *Psychological Review*, 114(1), 203-209. doi:10.1037/0033-295X.114.1.203
- Wixted, J.T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, 117(4), 1025-1054. doi:10.1037/a0020874
- Wixted, J.T., & Stretch, V. (2004). In defense of the signal-detection interpretation of Remember/Know judgments. *Psychonomic Bulletin & Review*, 11(4), 616-641.
- Wolk, D.A., Schacter, D.L., Lygizos, M., Sen, N.M., Holcomb, P.J., Daffner, K.R., & Budson, A.E. (2006). ERP correlates of recognition memory: Effects of retention interval and false alarms. *Brain Research*, 1096, 148-162. doi:10.1016/j.brainres.2006.04.050
- Wolk, D.A., Schacter, D.L., Lygizos, M., Sen, N.M., Chong, H., Holcomb, P.J., ... Budson, A.E. (2007). ERP correlates of Remember/ Know decisions: Association with the late posterior negativity. *Biological Psychology*, 75(2), 131-135. doi:10.1016/j.biopsycho.2007.01.005
- Woodmansee, J.J. (1966). Methodological problems in pupillographic experiments. In Proceedings from *The 74th Annual Convention of the American Psychological Association*, 1, 133-34.

- Woodmansee, J.J. (1967). The pupil reaction as an index of positive and negative affect. Paper presented at *The meeting of the American Psychological Association*. Washington, DC: American Psychological Association.
- Woodmansee, J.J. (1970). The pupil response as a measure of social attitudes. In C. Sumers (Ed.), *Attitude measurement*. Chicago: Rand-McNally.
- Wootton, A.C. (1910). *Chronicles of Pharmacy*. London: Macmillan & Co Ltd.
- Wright, P., & Kahneman, D. (1971). Evidence for alternative strategies of sentence retention. *Quarterly Journal of Experimental Psychology*, 23(2), 197-213. doi:10.1080/14640747108400240
- Wu, T.C.K., Allen, M.D., Goodrich-Hunsaker, N.J., Hopkins, R.O., & Bigler, E.D. (2010). Functional neuroimaging of symptom validity testing in traumatic brain injury. *Psychological Injury & Law*, 3(1), 50-62. doi:10.1007/s12207-010-9067-y
- Yamaji, K., Hirata, Y., & Usui, S. (2000). A method for monitoring autonomic nervous activity by pupillary flash response. *Systems & Computers in Japan*, 31(4), 22-31. doi:10.1002/(SICI)1520-684X(200004)31:4<22::AID-SCJ3>3.0.CO;2-W
- Yochim, B.P., Kane, K.D., Horning, S., & Pepin, R. (2010). Malingering or expected deficits? A case of herpes simplex encephalitis. *Neurocase*, 16(5), 451-460. doi:10.1080/13554791003623334
- Yonelinas, A.P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 20(6), 1341-1354. doi:10.1037/0278-7393.20.6.1341
- Yonelinas, A.P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, 25(6), 747-763. doi:10.3758/BF03211318
- Yonelinas, A.P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 25(6), 1415-1434. doi:10.1037/0278-7393.25.6.1415
- Yonelinas, A.P. (2001a). Components of episodic memory: The contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society B*, 356, 1363-1374. doi:10.1098/rstb.2001.0939

- Yonelinas, A.P. (2001b). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General*, 130, 361-379. doi:10.1037/0096-3445.130.3.361
- Yonelinas, A.P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, 46(3), 441-517. doi:10.1006/jmla.2002.2864
- Yonelinas, A.P., & Jacoby, L.L. (1996). Noncriterial recollection: Familiarity as automatic, irrelevant recollection. *Consciousness & Cognition*, 5(1-2), 131-141. doi:10.1006/ccog.1996.0008
- Yonelinas, A.P., Kroll, N.E., Dobbins, I., Lazzara, M., & Knight, R.T. (1998). Recollection and familiarity deficits in amnesia: Convergence of remember-know, process dissociation, and receiver operating characteristic data. *Neuropsychology*, 12, 323-339. doi:10.1037/0894-4105.12.3.323
- Yoss, R.E. (1969). The sleepy driver: A test to measure ability to maintain alertness. *Mayo Clinic Proceedings*, 44(11), 769-783.
- Yoss, R.E. (1970). The inheritance of diurnal sleepiness as measured by pupillography. *Mayo Clinic Proceedings*, 45(6), 426-437.
- Yoss, R.E., Moyer, N.J., Carter, E.T., & Evans, W.E. (1970). Commercial airline pilot and his ability to remain alert. *Aerospace Medicine*, 41, 1339-1346.
- Yoss, R.E., Moyer, N.J., & Hollenhorst, R.W. (1970). Pupil-size and spontaneous pupillary waves associated with alertness, drowsiness, and sleep. *Neurology*, 20(6), 545-554.
- Yoss, R.E., Moyer, N.J., & Ogle, K.N. (1969). The pupillogram and narcolepsy. A method to measure decreased levels of wakefulness. *Neurology*, 19(10), 921-928.
- Young, F.A., & Biersdorf, W. (1954). Pupillary contraction and dilation in light and darkness. *Journal of Comparative & Physiological Psychology*, 47(3), 264-268. doi:10.1037/h0057928
- Yovel, G., & Paller, K.A. (2004). The neural basis of the butcher-on-the-bus phenomenon: When a face seems familiar but is not remembered. *NeuroImage*, 21, 789-800. doi:10.1016/j.neuroimage.2003.09.034
- Zahn, T.P., Little, B.C., & Wender, P.H. (1978). Pupillary and heart rate reactivity in children with minimal brain dysfunction. *Journal of Abnormal Child Psychology*, 6(1), 135-147. doi:10.1007/BF00915789

- Zekveld, A.A., Kramer, S.E., & Festen, J.M. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear & Hearing, 32*(4), 498-510.
doi:10.1097/AUD.0b013e31820512bb
- Zhu, Y., Iba, M., Rajkowski, J., & Aston-Jones, G. (2004). Projection from the orbitofrontal cortex to the locus coeruleus in monkeys revealed by anterograde tracing. *Society for Neuroscience Abstracts, 30*, 211-213.
- Zuckerman, M. (1971). Physiological measures of sexual arousal in the human. *Psychological Bulletin, 75*(5), 297-329. doi:10.1037/h0030923

9. Appendices

A. Appendix: Experiments 1 and 1b

Study List

ANIMAL	COTTON	LETTER	PRINCE
ANSWER	DEGREE	MARBLE	PRISON
ARTIST	DRIVER	NATION	REPORT
BATTLE	ENGINE	NATURE	SADDLE
BRIDGE	FAMILY	ORANGE	SQUARE
BUDGET	FIGURE	PENCIL	STRING
CHURCH	FRIEND	PERSON	SYMBOL
CIRCLE	GENIUS	PLANET	THROAT
COLUMN	HEAVEN	POCKET	TONGUE
CORNER	ISLAND	POWDER	WINDOW

Recognition List

AVENUE	DOCTOR	MIRROR	SHADOW
BARREL	FATHER	MOTHER	SIGNAL
BEAUTY	FLOWER	OFFICE	SPIRIT
BOTTLE	GROUND	OXYGEN	STABLE
BRANCH	HUNTER	PALACE	STREET
BUTTER	JACKET	PARADE	SUMMER
CAREER	LAWYER	PERMIT	TEMPLE
CATTLE	LESSON	RECORD	VACUUM
COFFEE	LIQUID	RESULT	VALLEY
DANCER	MARKET	SCHOOL	WEAPON

B. Appendix: Experiment 2

Study List 1

BURIAL	LUXURY	RUDDER	SUPPLY
CHISEL	BELIEF	TRANCE	COLUMN
TRIPOD	COMEDY	SLEEVE	REPORT
AVENUE	PLEDGE	BLONDE	PARCEL
BUBBLE	VOLUME	REMARK	RIBBON
STRIPE	CAREER	TYRANT	BUTTON
SCHOOL	DINNER	COUSIN	SYMBOL
REWARD	RECALL	AUTUMN	HORROR
RUMBLE	KNIGHT	FUTURE	PLUNGE
SQUIRE	INFANT	NICKEL	COLLAR

Recognition List 1

LAWYER	MARVEL	COFFIN	BEAUTY
BASKET	CLOVER	PRINCE	SADDLE
MOTIVE	VESSEL	REASON	NECTAR
REMEDY	GARLIC	KITTEN	SPIDER
MOMENT	OBJECT	MASTER	CAVERN
AUTHOR	PERSON	PICKLE	POCKET
PRIEST	AMOUNT	WINDOW	BREEZE
PALACE	COTTON	ISLAND	WEIGHT
NATION	SHOVEL	CANNON	DESIGN
HAZARD	RATTLE	HONOUR	ARTIST

Study List 2

DENIAL	BREATH	MATTER	COPPER
FABRIC	RECORD	COURSE	THRONG
MARKET	SHAKER	BORDER	RESULT
SKETCH	SHRIEK	ANSWER	BRONZE
ANCHOR	GROWTH	BOTTLE	POWDER
BALLOT	MOTION	GASKET	FELLOW
BRANCH	DANCER	SUNSET	MINUTE
SALARY	NEPHEW	SPIRIT	MAGNET
SORROW	GRUDGE	SAFETY	GROUND
RUBBER	ORIGIN	RABBIT	STABLE

Recognition List 2

TUMBLE	GROCER	OYSTER	WINTER
PRAISE	APPEAL	SISTER	SPONGE
MORTAL	DECREE	REGRET	GIRDLE
PARADE	NUMBER	LAMENT	PIGEON
EMPIRE	SQUEAK	MOTHER	FACTOR
SPEECH	WEALTH	FOREST	ORANGE
MAIDEN	DAMAGE	INSECT	CHERRY
SEASON	CARROT	WISDOM	ERRAND
CINDER	DANGER	NAPKIN	BRANDY
DEBATE	GALAXY	CATTLE	THRILL

Study List 3

CELLAR	UMPIRE	PRISON	HATRED
NUTMEG	NATURE	SIGNAL	BLOUSE
PASTOR	CIRCUS	CORNER	JUNIOR
VANITY	ENGINE	SPHERE	FELINE
POETRY	NEEDLE	POSTER	RESORT

KETTLE
FAMILY
ESTATE
ENAMEL
SCREAM

VELVET
DEVICE
CHURCH
STROKE
THREAT

TEMPER
ROCKET
IMPACT
MUTTON
PEOPLE

REPAIR
HAMLET
WIZARD
VIRTUE
INCOME

Recognition List 3

TRAVEL
SYSTEM
MENACE
BISHOP
RELIEF
STARCH
SUPPER
TALENT
TURTLE
RACKET

RITUAL
HUNGER
HOCKEY
MEADOW
GUTTER
ARMOUR
TEMPLE
GINGER
CIRCLE
LENGTH

GRAVEL
GENIUS
CHAPEL
SPRING
METHOD
SHIVER
STRIDE
LOCKER
CEREAL
TENNIS

PARISH
HELMET
DELUGE
MIRROR
MANURE
LABOUR
MONKEY
STREET
WALLET
EXCUSE

Study List 4

GALLON
LESSON
SQUARE
BANKER
PHRASE
BUDGET
MARGIN
WALNUT
APATHY
VIOLIN

TIMBER
PEPPER
SATIRE
NARROW
ADVICE
LUMBER
EFFECT
HAMMER
IMPORT
TREATY

INJURY
CHILLY
AERIAL
POLLEN
CANDLE
GOSPEL
BULLET
DEFEAT
CRADLE
PILLOW

PLIERS
VICTIM
PROFIT
LEADER
NATIVE
PUZZLE
THEORY
CEMENT
GENDER
TICKET

Recognition List 4

WICKET
CUSTOM
RIDDLE
MISERY
DEGREE
CHROME
MEMBER
BARREL
REVOLT
BUCKET

WALRUS
HURDLE
FIGURE
TREMOR
SHOWER
COFFEE
HEALTH
KENNEL
CARPET
SULTAN

DECEIT
PATENT
SUMMER
THROAT
BUTTER
TOMATO
SECOND
ESCAPE
CHANCE
DRIVER

SEQUEL
TONGUE
SQUINT
BRIDGE
WEAPON
POTATO
JACKET
PRAYER
MEMORY
RESCUE

Calibration List

HUNTER	DOCTOR	SILVER	PENCIL
BANANA	STRAIN	PUBLIC	PLANET
ANIMAL	BURROW	WILLOW	THREAD
TERROR	FLOWER	LIQUID	CRISIS
OFFICE	PUDDLE	DOLLAR	MORTAR
NOTICE	BOTHER	TUNNEL	HUMOUR
CORPSE	PERIOD	HEIGHT	SHIELD
PESTLE	AGENCY	VALLEY	SHADOW
SEARCH	SHRIMP	INSULT	RATION
STATUE	FATHER	FINISH	PARDON

C. Appendix: Experiments 3 and 4

Word List A

OPINION	PERFECT	DESPAIR	BOATING
BALANCE	SHELTER	DISPUTE	COMFORT
DELIGHT	PRODUCE	MYSTERY	FAILURE
CONTACT	EXHIBIT	PASSAGE	PASSION
INSIGHT	PRODUCT	VARIETY	OPENING
TRIBUTE	EMOTION	ARTICLE	BENEFIT
SUCCESS	VICTORY	CONTEXT	
VEHICLE	ROUTINE	NEUTRAL	

Word List B

WITNESS	CONCERN	PARTNER	COMMAND
REALITY	LIBRARY	PROMISE	ATTEMPT
CIRCUIT	VETERAN	MESSAGE	CONCEPT
REMOVAL	TRIUMPH	EXTREME	OUTCOME
MEASURE	SILENCE	UNIFORM	PRIMARY
PAYMENT	HARMONY	LOYALTY	ABILITY
DIGNITY	WELCOME	EXPRESS	
EDITION	ANXIETY	ARRIVAL	

Grammar Study List A

XMMXM	XXRTTVM	VVTRTTVM	VVTRVTRVM
VTTVTM	VTTTTVM	XXRRTVTM	XXRRTTTVM
XXMRVM	XMMXRTVM	XXMRVTRVM	XMMXRTVTM
VVTRTVM	XMMMVRVM	XXRVTRTVM	

Grammar Study List B

XMTRM	XMVTRXM	XMVRMTRM	VVTTTTRXM
VVRMTM	VVTRXRM	VVTTRXRM	VVTTTRMTM
XMTRRM	VVTRXRRM	VVRMVRXRM	VTRRRRRRM
VTRRRRM	XMTRRRRM	XXRRRRRRM	

Grammar Recognition List A

VTVTM	VVTRVTM	XMMMMMXM	XMMMMMMXM
XXRVM	XXRTVTM	VTTTTVTM	VTTTVTRVM
XMMMVM	XXRVTM	XXRVTRVM	VVTRTTVTM
XXRTVM	XMMXRV	VTVTRVTM	VTTTTTTVM
XXRVTM	XXRRTTVM	VTTTTTTVM	XXRTTTTVM
VVTRVM	VTTVTRVM	VTTVTRTVM	XXRTVTRVM
VTTTVTM	VTVTRTVM	XMMXRRTVM	
XMMMMXM	VTTTTTVM	VTVTRTVM	

Grammar Recognition List B

VTRRM	XMVRMTM	VVTRMTRM	VVTRMVRXM
XXRRM	XXRRRRM	XMVTRMTM	VVTRXRRRM
VVRXRM	VVRMTRM	VVRMTRRM	VVTRMTRRM
VVTRXM	VVRXRRM	VVTTRMTM	XMVTTTTRXM
XMVRXM	XXRRRRRM	XMVTRMTRM	VVTTRXRRM
XXRRRM	VVTTTTRXM	VVTTRXRM	VVRXRRRRM
XMVRXRM	VTRRRRRM	XMVRMVRXM	
VVTTRXM	XMVTRXRM	XMVRMTRRM	

D. Appendix: Experiment 5

Study List 1

UNIFORM	STATION	EDITION	GRAMMAR
BLISTER	TROUBLE	MUSTARD	LOYALTY
ANTIQUE	CITIZEN	TRACTOR	FORTUNE
SKYLARK	MERCURY	MIRACLE	PATIENT
DEPOSIT	DIAMOND	SPEAKER	MINERAL
POVERTY	LEATHER	WITNESS	VEHICLE
SUCCESS	MONSOON	DUNGEON	PALETTE
LEAFLET	PRODUCT	WARRIOR	EMERALD
OUTCOME	ESSENCE	BRAVERY	FEELING
COMPANY	DREAMER	MANSION	PROBLEM

Recognition List 1

COLONEL	CAPSULE	VILLAGE	HADDOCK
PAYMENT	GARMENT	PATTERN	EMOTION
STUDENT	RECITAL	IMPULSE	PROTEST
HONESTY	DOORWAY	STADIUM	CHANNEL
BALANCE	MACHINE	COTTAGE	INSIGHT
BLOSSOM	RECEIPT	RACQUET	SOCIETY
PIONEER	BOREDOM	PLATTER	EVENING
CABINET	FORFEIT	CONTROL	ABILITY
BANDAGE	SOLDIER	CHARITY	BEDROOM
TRUMPET	TRIBUTE	STEEPLE	HOLIDAY

Study List 2

MERMAID	PEASANT	RECRUIT	TOURIST
TYPHOON	PARTNER	WEATHER	ARRIVAL
COSTUME	SILENCE	CHUCKLE	MESSAGE
HERRING	EMPEROR	BLANKET	COUNCIL
THUNDER	TROLLEY	FIGMENT	SEAWEED
GALLERY	LAUNDRY	DENTIST	MORNING
JOURNAL	FANTASY	LOBSTER	INQUIRY
HARVEST	GLITTER	FLANNEL	CANTEEN
SLIPPER	PORTION	MILEAGE	QUARTER
DIGNITY	RESIDUE	SUBJECT	CUISINE

Recognition List 2

VISITOR	PASTURE	ABDOMEN	DUCHESS
BUILDER	CENTURY	ARTICLE	COLLEGE
SHUTTER	BOATING	MIXTURE	MONARCH
SULPHUR	FREEDOM	CAPTAIN	CULTURE
SETTLER	TRAGEDY	CABBAGE	BROTHER
REGENCY	CARAVAN	CEILING	JUSTICE
SEGMENT	HARNESS	SERVICE	BUTCHER
PACKAGE	FOREARM	REALITY	FURNACE
HATCHET	PROFILE	CONCERT	PIANIST
APOLOGY	FASHION	APRICOT	WELCOME

Study List 3

RETREAT	THIMBLE	RESPECT	HISTORY
ASPIRIN	CREATOR	MALARIA	MEETING
GLACIER	TEMPEST	BOUQUET	FALLACY
REFEREE	OATMEAL	SPINACH	ROMANCE
LANTERN	THOUGHT	SUNBURN	VINEGAR

ACCOUNT	LIBRARY	PRELUDE	MYSTERY
GRAVITY	PURPOSE	OFFICER	TEACHER
SPANGLE	STEAMER	FACTORY	HYGIENE
HAIRPIN	ATHLETE	KNUCKLE	TOBACCO
IMPRINT	PHANTOM	VOLCANO	NURSERY

Recognition List 3

SESSION	COUNTRY	OPINION	SUSPECT
DYNASTY	DISEASE	HUSBAND	PASSAGE
CRYSTAL	REVENGE	BISCUIT	FLUTTER
PYRAMID	PICTURE	ECONOMY	GODDESS
TRAILER	PAINTER	TRIUMPH	BALLOON
COMMAND	KINGDOM	WHISTLE	FIELDER
STOMACH	FINANCE	MEASLES	HEROISM
BARGAIN	CHICKEN	ROBBERY	BAGPIPE
CIRCUIT	EPISODE	INTERIM	VICTORY
FAILURE	SLUMBER	SURFACE	SCHOLAR

Instructions

Scenario: We would like you to imagine that you are a person who has recently been involved in a car accident. You were unconscious for 15 minutes after the accident, and you had to spend one night in hospital for observation. Gradually your condition improved over the following months and you have now made a full recovery. Imagine that the purpose of the test that you are about to undertake is to determine whether the accident has produced any memory-impairments due to brain damage.

Instruction 1: Please perform to the best of your abilities on the recognition memory task, answering 'new' when you think a word is new and 'old' when you think a word is old.

Instruction 2: At your memory test, you decide to exaggerate the effects of your accident in case there is extra compensation money available. Please produce responses that would convince an examiner that you still have memory loss. Impairments should be presented in a "believable" manner, and major exaggerations, such as not remembering anything should be avoided, because even if you were performing at chance you would still get about 50% of the answers correct. £10 worth of book vouchers will be awarded to the individual who best manages to simulate a believable memory deficit.

Instruction 3: Please say 'new' to new words and also say 'new' to old words that you recognise.

E. Appendix: Experiment 6**Study List**

NUMBER	FAMILY	SYMBOL	CAREER
MARBLE	MIRROR	GROUND	BOTTLE
SPRING	SEASON	COUSIN	PERSON
SUPPER	CORNER	AUTUMN	OFFICE
PROFIT	REPORT	DANGER	SIGNAL
SILVER	BRIDGE	FIGURE	NATURE
MOTHER	DINNER	DANCER	ESTATE
TONGUE	CELLAR	HEIGHT	ORANGE
ANIMAL	PALACE	BORDER	SPEECH
PRISON	PENCIL	POCKET	BEAUTY

Recognition List

LIQUID	ANSWER	WEALTH	ARTIST
PLANET	BUDGET	COTTON	FRIEND
LAWYER	MARKET	DRIVER	POETRY
COLUMN	DOCTOR	BARREL	WINDOW
MINUTE	CHURCH	STREET	BUTTER
COFFEE	SHADOW	LESSON	LEADER
SADDLE	SISTER	POWDER	PARADE
FLOWER	LETTER	DEGREE	SUMMER
COMEDY	ISLAND	SCHOOL	PRAYER
RECORD	CENTER	CIRCLE	ENGINE

F. Appendix: Experiment 7**Study List**

BLANKET	SPINACH	FURNACE	LEATHER
DIAMOND	GALLERY	APOLOGY	PADLOCK
BALLOON	ATHLETE	BISCUIT	HAIRPIN
INCENSE	ECONOMY	LIGHTER	COTTAGE
FACTORY	CONCERT	ALGEBRA	MINERAL
KNUCKLE	BUTCHER	ACADEMY	CARAVAN
ANTIQUE	BEDROOM	SOLDIER	COSTUME
AMATEUR	ASPIRIN	EPISODE	MEASLES
BOUQUET	BROTHER	PYRAMID	BANDAGE
CHICKEN	LAUNDRY	CABBAGE	LANTERN

Recognition List

HARVEST	MUSTARD	ARCHERY	BOREDOM
---------	---------	---------	---------

GESTURE	PEASANT	CAPTAIN	CANTEEN
RAINBOW	LOBSTER	ALCOHOL	CABINET
FINANCE	FLANNEL	EMERALD	HOLIDAY
MANSION	BAGPIPE	GLITTER	DUNGEON
SLIPPER	SHUTTER	ROBBERY	SPEAKER
BLISTER	DISPLAY	JOURNAL	NURSERY
FASHION	CITIZEN	BUILDER	MISSILE
COURAGE	BLOSSOM	CEILING	CRYSTAL
HOSTAGE	BAYONET	FORTUNE	DENTIST

G. Appendix: Experiments 8 and 9

Weak Items – Block 1

SPLIT	BLOOM	GUIDE	SMART
CROWN	KNOCK	CREEK	SHEEP
ROCKY	REBEL	STEAK	ARISE
RAZOR	LOGIC	PITCH	NOBLE
CRIME	ALIEN	CHOSE	STORM
ALIKE	CHEEK	PUPIL	CIGAR
BLEAK	RIGID	BRAVE	
CABIN	SPRAY	STICK	

Strong Items – Block 1

MOUND	GUEST	CRAZY	CREAM
NURSE	UNITE	DRIFT	ADOPT
VOCAL	PLATE	PROSE	FAINT
HANDY	LOYAL	SHOCK	CRAWL
STAKE	PRINT	MOTEL	WIRED
FALSE	MORSE	ENTRY	YIELD
SMELL	WAIST	FRUIT	
THANK	SWIFT	GLOBE	

New Items – Block 1

LABEL	ORBIT	ANGEL	UPSET
BRIDE	GLORY	ELITE	CHARM
PATCH	PAINT	FAULT	BELLY
MAKER	ARGUE	CLOUD	CHAOS
BLADE	TOAST	SOLAR	FLOCK
CLUNG	TENSE	BLAST	AROSE
SLATE	BAKER	BENCH	
STEEP	STEAM	SPARK	

Weak Items – Block 2

ESSAY	SQUAD	ATLAS	MERCY
BRICK	RANCH	STOVE	SHAME
MOVIE	OCEAN	PENNY	BADLY
EXACT	DOUGH	TROOP	WOUND
BEARD	SHADE	SPADE	BACON
GLAZE	SUNNY	NERVE	CLERK
FLOOD	DUSTY	FOCUS	
FLUSH	TOWER	DRAIN	

Strong Items – Block 2

GUILT	DRILL	CLOCK	BATON
CHILL	INPUT	TALES	MEDIA
SEWER	CHASE	BOOST	ALARM
CRUDE	CYCLE	VAGUE	FOLLY
TWIST	BRAND	TRACE	RIDER
TIGHT	MAYOR	CURSE	DIRTY
RADAR	CIVIC	SWISS	
CREST	GRAVE	QUEST	

New Items – Block 3

SPORT	NOISE	CANDY	DODGE
BURST	MOUSE	ATTIC	DAIRY
SKIRT	COACH	ALERT	SUITE
MERGE	SAINT	STRAW	EAGER
POUND	HURRY	LYRIC	SLAVE
VIVID	GRAPH	DEBUT	GRACE
ADULT	OWNER	HABIT	
DEALT	TRACT	REACT	

Weak Items – Block 3

GHOST	ACTOR	PURSE	TOOTH
IMPLY	SLIDE	SAUCE	GLOOM
GROVE	HELLO	IVORY	SCOPE
LUNAR	FORGE	LEMON	GIANT
SUGAR	REALM	LEVER	BRACE
THUMB	HONEY	WIDTH	BUNCH
LAUGH	VERSE	IRONY	
QUOTE	ADMIT	GAUGE	

Strong Items – Block 3

SHIRT
DEVIL
NAVAL
ROAST
PANIC
MAGIC
FLASH
ELDER

STUFF
RALLY
GRAIN
CLIFF
PROOF
WRIST
FERRY
FLUID

TRICK
ARROW
PIANO
STUCK
LEASE
DELAY
GRILL
BLANK

CRASH
DITCH
JUICE
WORST
STARE
PAUSE

New Items – Block 3

FEVER
LOBBY
PASTE
ARRAY
TOKEN
FENCE
GRASP
ONION

RIVAL
MERIT
SLOPE
ALoud
CRAFT
VIRUS
HAVEN
DRAFT

BLAME
LAYER
STERN
AMPLE
AWFUL
AWAKE
CHEAP
CLIMB

MASON
BLOND
HARSH
TREAT
PRIZE
FLEET