PSYCHOLOGY MISCELLANY

No.41 - November 2012

Kevin Brewer

ISSN: 1754-2200

Orsett Psychological Services
PO Box 179
Grays
Essex
RM16 3EW
UK

orsettpsychologicalservices@phonecoop.coop

This document is produced under two principles:

1. All work is sourced to the original authors. The images are all available in the public domain (most from http://commons.wikimedia.org/wiki/Main_Page). You are free to use this document, but, please, quote the source (Kevin Brewer 2012) and do not claim it as you own work.

2. Details of the author are included so that the level of expertise of the writer can be assessed. This compares to documents which are not named and it is not possible to tell if the writer has any knowledge about their subject.

Kevin Brewer BSocSc, MSc

An independent academic psychologist, based in England, who has written extensively on different areas of psychology with an emphasis on the critical stance towards traditional ideas.

A complete listing of his writings at http://kmbpsychology.jottit.com.

CONTENTS

1. SCIENTIFIC PSYCHOLOGY AND THE PARANORMAL

1.1. INTRODUCTION

Scientific psychologists present their method as the way to establish truth about behaviour [1]. Science generally "may be understood to entail the precise, observer-independent measurement of phenomena that are progressively or cumulatively employed to evaluate falsifiable theoretical explications of the phenomena" (Gone 2011 p236). Thus the scientific method applied appropriately is unbiased, and produces universally applicable findings [2]. The critical approach to psychology challenges such claims of objectivity [3].

The scientific tradition is underpinned by Western philosophical traditions like rationalism (use of logic), empiricism (systematic observation of the world), and scepticism (subjective interpretations are questioned) (Gone 2011). Thus, in fact, science is "a historical product of the seismic cultural shifts wrought by the Western Enlightenment" (Gone 2011 p237).

Gone (2011) outlined the following limitations of "psychological science":

i) It is "never adopted or deployed outside of culturally constituted interests, objectives, and motivations, thereby requiring ongoing critical

[1] Gone (2011) defined "psychological science" as "the application of methods originally developed for the natural sciences to the study of what later came to be known as mind and behaviour" (p237).

[2] For example, Goette et al (2006) found that behaviour studied in laboratory experiments does not necessarily generalise to real-life. Swiss males undergoing compulsory military service were randomly assigned to groups which had little interaction (minimal groups; MG; typical of laboratory experiments) or a high degree of interaction (social groups; SG; typical of real-life) to play a sharing game. An individual could share their points or not with another player who was either an ingroup or outgroup member. Sharing (co-operation) was more common with ingroup members than with outgroup ones. Co-operation was greater in the SG than the MG, where the opportunity for interaction had created stronger bonds within the ingroup.

[3] "One problem in defining science is that it is difficult to generalise from palaeontology to particle physics or from astronomy to economics in order to distil any common activities out of the variety of recognised sciences. In fact, philosophers of science have demonstrated that there is no method or set of procedures shared by all sciences or scientists; some have even decried the emphasis on method itself as constraining to scientific progress..." (Gone 2011 p236).

engagement with the overt and covert agendas served by its findings" (p239).

ii) An inherent bias that favours publication of positive and groundbreaking findings over negative or confirming ones.

iii) It "depends on faith in its method for confidence in its findings. Once findings extend too far beyond preconceived notions of truth, however, psychologists – again, like scientists more generally – will dismiss such findings on the presumption of faulty methods and often declare such work as beyond the bounds of disciplinary inquiry" (p240).

iv) The problem of applying methods from natural science to the study of the mind, which can lead to "sterile results and impoverished knowledge" (p240).

Lehrer (2010) talked about the "decline effect" in research where subsequent replications of a study find a smaller effect. For example, Ioannidis (2005) found that a number of "established facts" about health had subsequently been contradicted or had their effect size downgraded (eg: daily low-dose aspirin to prevent heart attack and stroke) [4]. "The decline effect is troubling because it reminds us how difficult it is to prove anything" (Lehrer 2010).

One reason for the "decline effect" may be "publication bias". Initially, only studies with positive results are accepted for publication, but in time this changes to a preference for publishing studies with negative findings.

---

[4] Ioannidis (2005) examined 45 studies that produced key findings in medical research between 1990 and 2003, and found that seven were contradicted by subsequent studies, while twenty were replicated. Eleven studies were "unchallenged", and the remainder were found to have stronger effects in subsequent studies. Ioannidis (2005) observed that: "The examination of contradictions and refutations offers a fascinating look at the process of science. Four of the highly cited articles examined herein were refuting investigations with 'negative' results. However, in a sense, even the other highly cited articles with 'positive' results refuted prior knowledge and practice by introducing new concepts and proposing new interventions. We should acknowledge that there is no proof that the subsequent studies and meta-analyses were necessarily correct. A perfect gold standard is not possible in clinical research, so we can only interpret results of studies relative to other studies. Whenever new research fails to replicate early claims for efficacy or suggests that efficacy is more limited than previously thought, it is not necessary that the original studies were totally wrong and the newer ones are correct simply because they are larger or better controlled. Alternative explanations for these discrepancies may include differences in disease spectrum, eligibility criteria, or the use of concomitant interventions. Different studies on the same question are typically not replicas of each other. In fact discrepancies may be interesting on their own because they require careful scrutiny of the data and reappraisal of our beliefs. Thus, it is probably not surprising that the citation rate of these refuted studies did not seem to be much affected. Nevertheless, the controversy generates considerable uncertainty for clinical practice and none of the contradicted interventions is currently recommended by practice guidelines" (p227).

But the "decline effect" has been seen in unpublished studies. In this case, methodological weaknesses may be responsible for differences in findings (eg: no blinding), and researchers engaging in "significance chasing" ("finding ways to interpret the data so that it passes the statistical test of significance"; Lehrer 2010).

1.2. THE PARANORMAL

The scientific method is seen as the best approach to assessing paranormal (psi) claims as covered in parapsychology. Claims of the ability to foretell the future, for example, "eventually are either accounted for by normal means, or else they disappear under controlled conditions" (Shermer 2011).

Recently, however, Daryl Bem has reported findings from experimental work which support pre-cognition (Bem 2011; appendix 1A). In the most reported of the nine experiments, participants had to predict on which of two screens a picture would appear before the computer randomly chose one. The pictures used were either neutral, negative or erotic. Only for the erotic pictures the participants were correct statistically significantly more than chance (53.1% vs 50% chance).

Shermer (2011) criticised the findings as evidence supporting pre-cognition [5]. He argued that there may not be a natural means to explain the results, but that is not the same as evidence for pre-cognition. For example, it is possible that the findings are statistically significant "just by chance" (appendix 1B).

Replications have failed to confirm the findings - Galak et al (2012) experiments 8 and 9 [6], and Ritchie et al (2012) experiment 9 (Bhattacharjee 2012).

Moulton and Kosslyn (2008) observed: "If psi exists, sciences from physics to psychology may require fundamental revision. If psi does not exist, roughly half of the general population... should be disabused of their fallacious beliefs" (p182). This was the logic behind their study of telepathy with participants in functional magnetic resonance imaging (fMRI) scanners. The assumption being that brain activity would occur during telepathy. Nineteen pairs of individuals were divided into "senders" and "receivers". The task of the sender in

[5] Gone (2011) observed a certain irony that "psychological science" emphasises the importance of rigorous methods to establish truth, and then criticises parapsychological research that uses such methods when the findings appear statistically significant.

[6] Galak, J et al (2012) Correcting the past: Failures to replicate psi (unpublished paper at http://ssm.com/abstract=2001721).

one room was to telepathically communicate a picture randomly presented to them, while the receiver in the fMRI scan had to choose between two pictures offered.

Out of 3867 trials, the receivers correctly guessed the picture being sent on 50% of occasions (ie: same as chance). The fMRI recorded no difference in brain activity in the telepathic or control trials. Moulton and Kosslyn (2008) were definite in their conclusion:

> The results support the null hypothesis that psi does not exist. The brains of our participants — as a group and individually — reacted to psi and non-psi stimuli in a statistically indistinguishable manner. Given the relatively large number of participants, the use of fixed-effects statistics, the extensive activation elicited separately by both types of stimuli, the subtle psychological effects revealed in the much smaller data set from a single participant, and the non-psi effects we documented on a group level using identical statistical criteria, a lack of statistical power does not reasonably explain our results. Even if the psi effect were very transient, as are many mental events, it should have left a footprint that could be detected by fMRI — as did the other subtle effects we detected. In particular, the large and massively significant activation revealed by our arousal contrast shows that that the psi effect, if it exists, must be substantially smaller than the effect of arousal on brain activity (p189).

1.3. NULL RITUAL

Gigerenzer et al (2004) highlighted the problems for psychology in following the "null ritual" (or "null hypothesis significance testing procedure"; NHSTP; Chow 1998) - "(1) Set up a statistical null hypothesis of 'no mean difference' or 'zero correlation'. Don't specify the predictions of your research hypothesis or of any alternative substantive hypotheses.
(2) Use 5% as a convention for rejecting the null. If significant, accept your research hypothesis.
(3) Always perform this procedure".

When the means of the experimental group and control group are statistically significantly different (p-value = 0.05), researchers can interpret this finding incorrectly in relation to the null hypothesis. Such a difference does not (Gigerenzer et al 2004):

- Absolutely disprove the null hypothesis nor absolutely prove the experimental hypothesis. "Significance tests provide probabilities, not definite proofs" (Gigerenzer et al 2004).

- Tell the probability of the null hypothesis being true nor the probability of the experimental hypothesis

being true.

- Tell "if you decide to reject the null hypothesis, the probability that you are making the wrong decision".

- Tell that if the experiment was repeated many times, there would be a significant result on 95% of occasions.

There is a tendency to overestimate what can be concluded from a p-value. A high percentage of students and lecturers in psychology with experience of statistics tend to endorse at least one of the statements above (over 80%; Gigerenzer et al 2004) [7].

The p-value is the probability of the observed data given that the null hypothesis is true (Gigerenzer et al 2004). The smaller the p-value (or more significant the difference between the group means, say) the less likely the observed data is by chance.

Gigerenzer et al (2004) outlined their concerns:

a) Having only a null hypothesis to disprove is limiting. "Hypotheses testing should always be competitive; that is, the predictions of several hypotheses should be specified" (Gigerenzer et al 2004).

b) Talking about rejecting or accepting hypotheses is unhelpful.

c) Using a conventional level of significance like 5% (0.05) is just that - convention. Gigerenzer et al recommended reporting the exact level of significance (eg: p = 0.011 or 0.053).

d) The p-value conveys a limited amount of information. "Thus, report p-values together with information about effect sizes, or power, or confidence intervals" (Gigerenzer et al 2004) [8].

1.4. APPENDIX 1A - BEM'S EXPERIMENTS

Bem (2011) distinguished precognition (conscious

---

[7] Tversky and Kahneman (1971) asked the following question: "Suppose you have run an experiment on 20 subjects and have obtained a significant result which confirms your theory (...p<0.05, two-tailed). You now have cause to run an additional group of 10 subjects. What do you think the probability is that the results will be significant, by a one-tailed test, separately for this group?" (quoted in Bem 2011). The majority of psychologists asked said over 50% (0.50), when the correct answer is approximately 48% (0.48) (Bem 2011).

[8] Rosenthal (1990) observed: "Given the levels of statistical power at which we normally operate, we have no right to expect the proportion of significant results that we typically do expect, even if in nature there is a very real and very important effect" (quoted in Bem 2011).

cognitive awareness) and premonition (affective apprehension) as "special cases" of "the anomalous retroactive influence of some future event on an individual's current responses, whether these responses are conscious or non-conscious, cognitive or affective" (p407). Bem tested this idea of "time-reversing" in nine experiments on four variations of the phenomena.

1. Two experiments on precognitive approach and avoidance.

Bem (2011) suggested that "our physiology can anticipate unpredictable erotic or negative stimuli before they occur. Such anticipation would be evolutionarily advantageous for reproduction and survival if the organism could act instrumentally to approach erotic stimuli and avoid negative stimuli" (p408).

Experiment 1

One hundred psychology students at Cornell University, USA, had to guess which curtain on a computer screen of two had a picture behind it. Over 36 trials different types of picture were presented - erotic (couples engaged in sexual acts), non-erotic romantic, negative, positive, and neutral. The participants made their guess, and then the computer randomly chose the curtain. Thus this was a test of detecting a future event (precognition).

Chance is 50% correct, and the hit rate for the erotic pictures was 53.1% ($p = 0.01$) and 49.8% for the other pictures (49.6% for neutral pictures, 51.3% negative pictures, 49.4% positive pictures, and 50.2% non-erotic romantic).

Individuals who self-rated as extravert before the experiment had a hit rate of 57.6% for the erotic pictures ($p = 0.00002$).

Experiment 2

One hundred psychology undergraduates at Cornell University were asked to record a preference for one of two identical pictures on a computer screen. After their choice the computer would randomly choose to show negative subliminally images through one picture. This experiment was testing the ability to predict which picture would have the negative images and to choose the opposite one.

Participants chose the picture not associated with future negative images on 51.7% of trials ($p = 0.009$), and extraverts on 53.5% of trials ($p = 0.002$).

2. Two experiments of affective priming.

Affective priming is where an individual is shown a picture and must judge as quickly as possible if it is pleasant or unpleasant while subliminally (below conscious perception) a positive or negative word (prime) is flashed on the screen. If the word and the picture are congruent (eg: positive word and pleasant picture) individuals' reaction times to decide is quicker than if the word and picture are incongruent (eg: positive word and unpleasant picture). This is called the contrast effect.

Experiment 3

This was an experiment using retroactive priming (ie: the prime is flashed after the decision was made). One hundred more students were shown 64 pictures and asked to press a different computer key if it was pleasant or unpleasant. Half the trials were standard priming and half were retroactive priming (figure 1.1).

Participants were 23.6 ms faster on average on the congruent than incongruent trials of the standard priming task ($p<0.00001$), and 15.0 ms faster on congruent trials of the retroactive priming task ($p =0.006$). Or put another way, 64.9% of participants were faster on congruent than incongruent trials of the standard version ($p = 0.002$) and 60.8% on the retroactive version ($p = 0.021$).

(Based on Bem 2011 figure 2 p414).

Figure 1.1 - Standard and retroactive priming tasks.

Experiment 4

This was a replication of experiment 3, but the stimuli used were more specific - eg: picture of fruit (positive prime = "luscious", negative prime = "bitter"), or picture of menacing pit bull (positive prime = "friendly", negative prime = "threatening"). In this experiment with 100 psychology students, the mean reaction time was 27.4 ms faster to decide if the picture was pleasant or unpleasant on congruent than incongruent

trials of the standard priming task (p = <0.00001) and 16.5 ms on the retroactive version (p = 0.023).

3. Three experiments on habituation.

In the standard habituation task, a participant is shown a picture subliminally a few times before being given a choice of two pictures. There is a tendency to prefer the picture presented subliminally though the participant has no perception of seeing it before.

Experiment 5

In the retroactive habituation task (figure 1.2), participants make a preference for one of the two pictures, and then the computer randomly decides which one to present subliminally. This is a test of the ability to predict which picture will be shown subliminally. The pictures were either a pair of negatively arousing ones or neutral.

Among the one hundred psychology students, there was a preference for the picture that would be shown subliminally afterwards (hit) on 53.1% of trials with the negative pictures (p = 0.014) (women 53.6%, p = 0.014; men 52.45, not significant), but only 49.4% for the neutral pair of pictures.

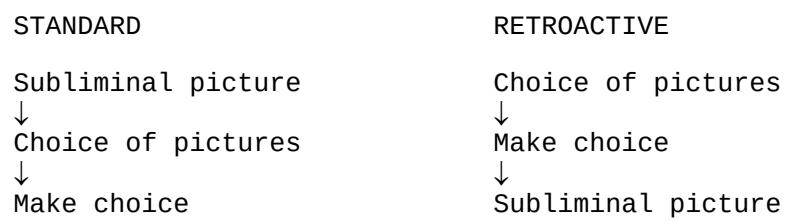


Figure 1.2 - Standard and retroactive versions of habituation task.

Experiment 6

This was a replication of experiment 5 with the addition of pairs of erotic pictures as stimuli.

On trials with negative pictures, the hit rate was 51.8% (p = 0.037), and 49.3% for neutral pictures (not significant), but 48.2% for the erotic pictures. This was a significant preference (p = 0.041) for the picture that would not be presented subliminally afterwards. Bem explained this finding as a product of repeated exposure leading to boredom and less preference.

Experiment 7

In this experiment, Bem sought to show that retroactive boredom could be produced by subliminally presenting a picture many times. After choosing between two pictures, 200 psychology students were randomly shown one picture subliminally ten times (as opposed to 4 or 6 times in previous experiments). It was predicted that this time due to boredom, participants would choose the picture that would not be shown subliminally.

There was only a significant preference among extraverts for the picture not to be shown (52.1% of trials; p = 0.019).

4. Two experiments on facilitation of recall.

Recall can be improved if participants are given cues after learning the list of words and before being asked to remember them. If individuals are given cues for some of the words on the list, their recall of these will be better than of the words not cued.

Experiment 8

In this experiment, participants were shown a list of 48 common nouns and then asked to recall them. Then after recall, the participants were given cues related to 24 words. Retroactive facilitation (or precognition) would mean more recall of the 24 words to be cued than the 24 not subsequently cued (figure 1.3).

Significantly more words to be cued were recalled than the words not to be cued (2.27% more; p = 0.029) by the 100 students.

Figure 1.3 - Standard and retroactive versions of facilitation of recall.

Experiment 9

In a replication of experiment 8, 4.21% more words to be cued were recalled than words not to be cued (p =

0.002) [9].

1.4.1. Bem Debate

Wagenmakers et al (2011) observed: "Instead of revising our beliefs regarding psi, Bem's (2011) research should instead cause us to revise our beliefs on methodology: the field of psychology currently uses methodological and statistical strategies that are too weak, too malleable, and offer far too many opportunities for researchers to befuddle themselves and their peers" (p426).

Wagenmakers et al (2011) criticised Bem (2011) for three mistakes in statistical analysis of the data [10]:

i) Confusion between explanatory and confirmatory studies - Confirmatory studies set out to test a hypothesis (ie: to confirm or disconfirm it), while exploratory studies are "fishing expeditions" that look for patterns in the collected data (or "torture the data until they confess"; Wagenmakers et al 2011).

Wagenmakers et al (2011) pointed out: "As such, there is nothing wrong with fishing expeditions. But it is vital to indicate clearly and unambiguously which results are obtained by fishing expeditions and which results are obtained by conventional confirmatory procedures. In particular, when results from fishing expeditions are analysed and presented as if they had been obtained in a confirmatory fashion, the researcher is hiding the fact that the same data were used twice: first to discover a new hypothesis, and then to test that hypothesis. If the researcher fails to state that the data have been so used, this practice is at odds with the basic ideas that underlie scientific methodology" (p427). Bem was accused of explanatory analysis of his data in experiments 1, 3, 5, and 6.

Bem et al (2011) argued that the studies were not exploratory. For example, in experiment 1, the hypothesis about erotic images was unambiguously directional. Bem (2011) used multiple statistical tests to analyse the data, not because the studies were exploratory, but to confirm that the conclusions held across different

---

[9] Wiseman (2010; quoted in Ritchie et al 2012) noted that in experiments 8 and 9 when participants misspelt words and Bem checked them, he was not blind to whether the word was to be cue or not. This opened the possibility of unconscious bias.

[10] "We realise that the above flaws are not unique to the experiments reported by Bem (2011). Indeed, many studies in experimental psychology suffer from the same mistakes. However, this state of affairs does not exonerate the Bem (2011) experiments. Instead, these experiments highlight the relative ease with which an inventive researcher can produce significant results even when the null hypothesis is true" (Wagenmakers et al 2011 p427).

statistical tests of the data (Bem et al 2011).

ii) There was "insufficient attention to the fact that the probability of the data given the hypothesis does not equal the probability of the hypothesis given the data (ie: the fallacy of the transposed conditional)" (Wagenmakers et al 2011 p426).

For example, the probability of a person being dead given that they were hanged (probability of the data given a hypothesis) and the probability that a person was hanged given that they are dead (probability of the hypothesis given the data) (Wagenmakers et al 2011). The upshot is that "extraordinary claims require extraordinary evidence", and a hit rate of 53.1% in Experiment 1 is not large enough to support a claim of the existence of pre-cognition. "Thus, in order to convince scientific critics of an extravagant or controversial claim, one is required to pull out all the stops. Even when Bem's (2011) experiments had been confirmatory (which they were not...), and even if they would have conveyed strong statistical evidence for precognition (which they did not...), eight experiments are not enough to convince a sceptic that the known laws of nature have been bent. Or, more precisely, that these laws were bent only for erotic pictures, and only for participants who are extraverts" (Wagenmakers et al 2011 p429).

iii) The use of a statistical test that overstated the evidence against the null hypothesis - Wagenmakers et al (2011) recommended pitting an alternative hypothesis against a null hypothesis rather than just considering the null hypothesis alone. Using other statistical tests with this assumption, Wagenmakers et al found that Bem's data do not support a hypothesis of pre-cognition.

Bem et al (2011) responded that Wagenmakers et al (2011) had incorrectly characterised features of Bem's (2011) experiments which underestimated the support for pre-cognition.

Wagenmakers et al (2011) concluded that Bem's findings in favour of pre-cognition "may well be illusory" because of the statistical analysis used. "However, Bem (2011) played by the implicit rules that guide academic publishing — in fact, Bem (2011) presented many more studies than would usually be required. It would therefore be mistaken to interpret our assessment of the Bem (2011) experiments as an attack on research of unlikely phenomena; instead, our assessment suggests that something is deeply wrong with the way experimental psychologists design their studies and report their statistical results. It is a disturbing thought that many

experimental findings, proudly and confidently reported in the literature as real, might in fact be based on statistical tests that are explorative and bias... We hope the Bem (2011) article will become a signpost for change, a writing on the wall: psychologists must change the way they analyse their data" (Wagenmakers et al 2011 p431). Bem et al (2011) felt that the statistical knowledge required to use more sophisticated techniques is demanding a lot of psychologists as these techniques are not without problems.

Ritchie et al (2012) attempted to replicate Bem's (2011) experiment 9 at three separate laboratories in Britain with a total of 150 students. The participants were presented with 48 words to learn followed by a recall test. After this test, half of the words were cued (practice words) and half not (control words). Pre-cognition would produce recall of more of the practice than control words.

The words recalled were scored as Bem (2011) did, which was to subtract the number of control words from the number of practice words remembered for each participant (as converted into a percentage - weighted "differential recall percentage"; DR%) (table 1.1).

PRACTICE WORDS	CONTROL WORDS	CONCLUSION
6/24 = 25%	12/24 = 50%	DR% = -25 (evidence against pre-cognition)
12/24 = 50%	6/24 = 25%	DR% = +25 (evidence for pre-cognition)
12/24 = 50%	12/24 = 50%	DR% = 0 (no difference)

Table 1.1 - Examples of "differential recall percentage".

From the three replications Ritchie et al (2012) found a mean DR% score of -1.03 (ie: marginally more control than practice words recalled). Thus a failure to replicate Bem's (2011) findings despite following his procedure closely.

1.5. APPENDIX 1B - QUASI-EXPERIMENT

Example: Liu et al (2009)

Aim: To test if driving experience reduced the risk of an accident by increasing hazard perception among motorcycle riders using a motorcycle simulator.

Method: The participants were 49 riders in Melbourne, Australia, who responded to an advertisement at the Honda Rider Training Centre [11].

The riders were divided into four groups [12]:

- Experienced riders with a full car driving licence (ie: experienced in riding and driving) (EM-FD) (N = 12).
- Inexperienced riders with a full driving licence (ie: inexperienced rider but experienced driver) (IM-FD (N = 12).
- Novice riders with a full driving licence (NM-FD) (N = 15).
- Novice riders with a probationary driving licence (NM-PD) (N = 10) [13].

Participants used the Honda Riding Simulator (of a 400cc motorcycle) in three scenarios - rural ("Touring"), medium traffic ("Avenue"), and high traffic density ("Path"). Each scenario had eight hazards (eg: pedestrians stepping out into the road; a tight hair-pin turn with no warning sign beforehand; a truck stopped at the side of the road with its door opening suddenly).

The dependent variable was measured by number of "crashes", evaluation of hazards, and vehicle measures like speed.

Findings: The NM-PD group had significantly more crashes in all 3 scenarios than the EM-FD group. There was no difference in evaluation of hazards between the groups. Experienced riders approached hazards slower than novices.

Conclusion: Experienced riders performed better than novices.

1.6. REFERENCES

Bem, D.J (2011) Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect Journal of Personality and Social Psychology 100, 3, 407-425

Bem, D.J et al (2011) Reply: Must psychologists change the way they analyse their data? Journal of Personality and Social Psychology 101, 4, 716-719 [14]

[11] This is a volunteer sample.

[12] This study is a quasi-experiment because the allocation to the groups was pre-determined and randomly done by the experimenter. In other words, individuals could not have been in any of the groups, which is randomly allocation as in a "true" experiment.

[13] The quasi-independent variable is the amount of motorcycle riding experience.

[14] Draft version at http://dl.dropbox.com/u/8290411/ResponsetoWagenmaker.pdf.

Bhattacharjee, Y (2012) Paranormal psychologist Discover 3, 52-58

Chow, S.L (1998) Precis of "Statistical Significance: Rationale, Validity and Utility" Behavioural and Brain Sciences 21, 169-239

Gigerenzer, G et al (2004) The null ritual: What you always wanted to know about significance testing but were afraid to ask. In Kaplan, D (ed) The Sage Handbook of Quantitative Methodology for the Social Sciences Thousand Oaks, CA: Sage

Goette, L et al (2006) The impact of social ties on group interactions: Evidence from minimal groups and randomly assigned real groups American Economic Review 96, 2, 212-216

Gone, J.P (2011) Is psychological science a-cultural? Cultural Diversity and Ethnic Minority Psychology 17, 3, 934-942

Ioannidis, J.P.A (2005) Contradicted and initially stronger effects in highly cited clinical research Journal of American Medical Association 294, 2, 218-228

Lehrer, J (2010) The truth wears off New Yorker 13/12, 52-57 (Available at http://www.newyorker.com/reporting/2010/12/13/101213fa_fact_lehrer)

Liu, C.C et al (2009) Hazard perception abilities of experienced and novice motorcyclists: An interactive simulator experiment Transportation Research F 12, 325-334

Moulton, S.T & Kosslyn, S.M (2008) Using neuroimaging to resolve the psi debate Journal of Cognitive Neuroscience 20, 1, 182-192

Ritchie, S.J et al (2012) Failing the future: Three unsuccessful attempts to replicate Bem's "retroactive facilitation of recall" effect PLoS ONE 7, 3, e33423

Rosenthal, R (1990) Replication in behavioural research Journal of Social Behaviour and Personality 5, 1-30

Shermer, M (2011) Extrasensory pornception Scientific American May, p73

Tversky, A & Kahneman, D (1971) Belief in the law of small numbers Psychological Bulletin 2, 105-110

Wagenmakers, E-J et al (2011) Why psychologists must change the way they analyse their data: The case of psi: Comment on Bem (2011) Journal of Personality and Social Psychology 100, 3, 426-432

2. STUDYING ADHD IN THE 21ST CENTURY

2.1. STUDYING GENETICS

Attention-deficit hyperactivity disorder (ADHD) is characterised by hyperactivity, inattention, and impulsivity. It is reported in 2-5% of schoolchildren, with the vast majority of sufferers being male (Mill et al 2005).

It has been suggested that ADHD has a heritability of about 70%, and is due to numerous genes (Mill et al 2005). The focus of current research is to pin down the actual genes involved in the disorder and in each symptom. A number of different methods and techniques are used.

1. Dizygotic (DZ) (identical) twins

Example 1: Mill et al (2005)

329 pairs of male DZ twins from the Twins' Early Development Study (TEDS) (all twins born in England and Wales in 1994-6). Symptoms of ADHD were rated by the parent(s) at 2, 3, 4, and 7 years old, and by teachers at age 7. Evidence of a gene linked to dopamine (DAT1) associated with hyperactivity.

2. Family studies

Example 2: Stawicki et al (2006)

132 children from Michigan state, USA, with ADHD and 204 of their biological parents were compared to 78 controls and 131 of their parents for rates of ADHD among biological parents. Among parents of ADHD sufferers, 17% were diagnosed as having ADHD, which is significantly more than the 3% of control parents.

Example 3: Brookes et al (2007)

This study investigated 51 candidate genes linked to neurotransmitter regulation among 674 ADHD sufferers (probands) and their biological parents and siblings on the International Multi-centre ADHD Gene (IMAGE) project.

This involved twelve sites in eight European countries. Eighteen genes were found to be significantly more common in parents and probands than siblings, but it was not clear whether they are all associated with ADHD.

Example 4: Neale et al (2010)

This study reported a meta-analysis of three family studies - Children's Hospital of Philadelphia (CHOP), IMAGE, and the Pfizer-funded study from the University of California, Los Angeles, Washington University, and Massachusetts General Hospital (PUWMa). The focus was genome-wide association studies (which search the whole genome for common genes between parent and child ADHD sufferers) rather than candidate gene association studies (which concentrate on specific genes based on hypotheses). The researchers reported no significant associations between sufferers and biological parents as compared to controls and their parents.

3. Adoption studies

Example 5: Sprich et al (2000)

Twenty-five US children diagnosed with ADHD after adoption in the first year of life and 62 adopted first-degree relatives (parents, siblings) ("adopted ADHD" group); 101 non-adopted children with ADHD and 310 biological first-degree relatives ("biological ADHD" group); fifty children without ADHD and their 153 relatives ("controls"). The prevalence of ADHD in relatives was highest in the "biological ADHD" group suggesting the inheritance of the disorder (table 2.1).

GROUP	PREDICTION	PREVALENCE OF ADHD IN FIRST-DEGREE RELATIVES
Adopted ADHD	If ADHD high in relatives = support for environmental cause	parents: 6% siblings: 8%
Biological ADHD	If ADHD high in relative = support for genetic cause	parents: 18% siblings: 31%
Controls		parents: 3% siblings: 6%

(Source of data: Sprich et al 2000 figure 1 p1434)

Table 2.1 - Rates of ADHD in relatives of ADHD sufferers (probands).

2.1.1. Basic Genetics

Each cell in the body contains a complete set of chromosomes, which for humans is 46: 22 pairs the same for either sex and then XX or XY.

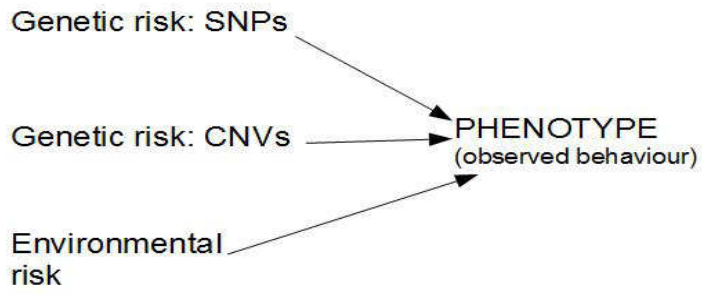
Chromosomes are made up of genes, which are based on DNA (deoxyribonucleic acid). DNA is made up of four bases: adenine (A), guanine (G), cytosine (C), and thymine (T). These are ordered around two chains wrapped together as the double helix.

The nature of DNA is such that A always pairs with T, and C and G. The sequence of bases is the genetic information.

Changes in the sequence of bases leads to changes in cell development. Point mutation is the substitution of one base for another, and is the most obvious example. Others include deletions (loss of sequence of bases), insertions (gaining of a piece of DNA), frameshift mutations (the loss of one base affecting the coding of others), and translocations (the breaking of part of a chromosome and reforming at a different site on a different chromosome).

Modern technology allows the genetic code of individuals to be compared for similarities and differences by looking at single-nucleotide polymorphisms (SNPs) (one gene pair) or copy number variations (CNV) (large number of gene pairs). The study of SNPs suggested that any two randomly selected human genomes vary by 0.1%, whereas the study of CNVs put the difference at 1% (Beckmann et al 2007).

The simple link between a gene (genotype) and an observed disease or behaviour (phenotype) is rare, mainly because multiple genes (polygenic) are involved in any phenotype. There is also the role of the environment (figure 2.1).



(Based on Beckmann et al 2007 figure 3 p643)

Figure 2.1 - Influences on the phenotype.

The link between genotype and phenotype is even more complex with the inclusion of an intermediate link called endophenotypes [15]. These are effects of genes that are not visible to the naked eye. For example, one version of a gene leads to a smaller head (visible trait), but the endophenotype produce by the gene is fewer cerebral neurons (which in turn leads to the smaller head).

Figure 2.2 gives an general example from psychiatry.

GENE	→	ENDOPHENOTYPE	→	PHENOTYPE
COMT*	→	P300 BRAINWAVE DIFFERENCES**	→	COGNITIVE DIFFERENCES IN SCHIZOPHRENIA***

* Codes catechol O-methyltransferase enzyme which inactivates catechols (eg: dopamine in prefrontal cortex) at post-synaptic sites. Specifically it codes for Valine (Val) or Methionine (Met) at position 158/108 on the genome (Flint and Munafo 2007).

** Measured by EEG around 300 milliseconds after participants respond to a random auditory or visual stimulus.

*** Eg: selective attention, working memory.

Figure 2.2 - Example of an endophenotype in schizophrenia.

2.2. BRAIN DIFFERENCES

The view on the brain of ADHD sufferers has moved from abnormalities in specific areas to dysfunction in connections between areas of the brain. Three aspects of brain connectivity are studied - anatomical (connections between structures in the brain), functional (eg: activity of the resting brain), and effective connectivity (influence of one brain region over another) (Konrad and Eickhoff 2010).

- Functional connectivity - Functional magnetic resonance imaging (fMRI) shows the resting state of the brain (known as the default-mode network; DMN). Studies have reported inconsistent findings about the DMN in ADHD sufferers as to whether it is over- (hyper) or under-active (hypo) compared to controls (Konrad and Eickhoff 2010).

- Structural connectivity - Diffusion tensor imaging (DTI) is used to study this. It shows the white matter pattern in the brain based on the diffusion of water molecules under magnetic force, which is restricted by the cell membranes and myelin sheath (Konrad and

[15] Singh and Rose (2009) advised cautious when searching for "biomarkers" in psychiatry because of the social, ethical, and legal problems inherent in their use.

Eickhoff 2010). Inconsistencies have been reported, for example, in white matter volume (either less or more than controls) (Konrad and Eickhoff 2010).

- Effective connectivity - Compared to controls differences have been found in the interaction of brain areas of ADHD sufferers during working memory tasks, for example, using fMRI (Konrad and Eickhoff 2010).

Konrad and Eickhoff (2010) emphasised the importance of studying the brain: "...it is necessary to learn how genes and environmental factors impact neural network architecture. Many of the early neurodevelopmental processes that are believed to be disrupted in ADHD are likely to be mediated by genetic mechanisms. Genetic and environmental factors are also believed to be involved in the continuity of the disorder, as well as in changes in ADHD symptomatology, throughout life" (p913).

2.3. REFERENCES

Beckmann, J.S et al (2007) Copy number variants and genetic traits: Closer to the resolution of phenotypic to genotypic variability Nature Reviews Genetics 8, 639-646

Brookes, K et al (2006) The analysis of 51 genes in DSM-IV combined type attention deficit hyperactivity disorder: Association signals in DRD4, DAT1 and 16 other genes Molecular Psychiatry 11, 934-953

Flint, J & Munafo, M.R (2007) The endophenotype concept in psychiatric genetics Psychological Medicine 37, 163-180

Kendler, R.S & Neale, M.C (2010) Endophenotype: A conceptual analysis Molecular Psychiatry 15, 8, 789-797

Konrad, K & Eickhoff, S.B (2010) Is the ADHD brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder Human Brain Mapping 31, 904-916

Mill, J et al (2005) Quantitative trait locus analysis of candidate gene alleles associated with attention deficit hyperactivity disorder (ADHD) in five genes: DRD4, DAT1, DRD5, SNAP-25, and 5HT1B American Journal of Medical Genetics, Part B (Neuropsychiatric Genetics) 133B, 68-73

Neale, B.M et al (2010) Meta-analysis of genome-wide association studies of attention deficit/hyperactivity disorder Journal of American Academy of Child and Adolescent Psychiatry 49, 9, 884-897

Singh, I & Rose, N (2009) Biomarkers in psychiatry Nature 460, 202-207

Sprich, S et al (2000) Adoptive and biological families of children and adolescents with ADHD Journal of American Academy of Child and Adolescent Psychiatry 39, 11, 1432-1437

Stawicki, J.A et al (2006) Family psychiatric history evidence on the nosological relations of DSM-IV ADHD combined and inattentive sub-types: New data and meta-analysis Journal of Child Psychology and Psychiatry 47, 9, 935-945

3. TWO DIAGRAMS FOR UNDERSTANDING BEHAVIOURS

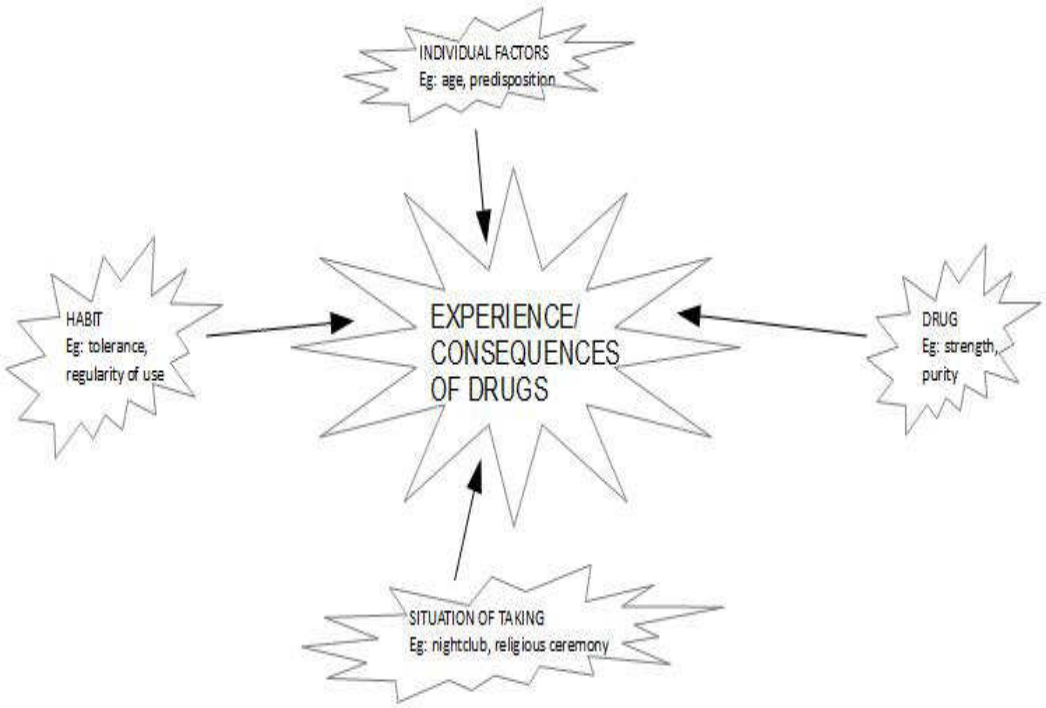


Figure 3.1 - Drug-taking experience.

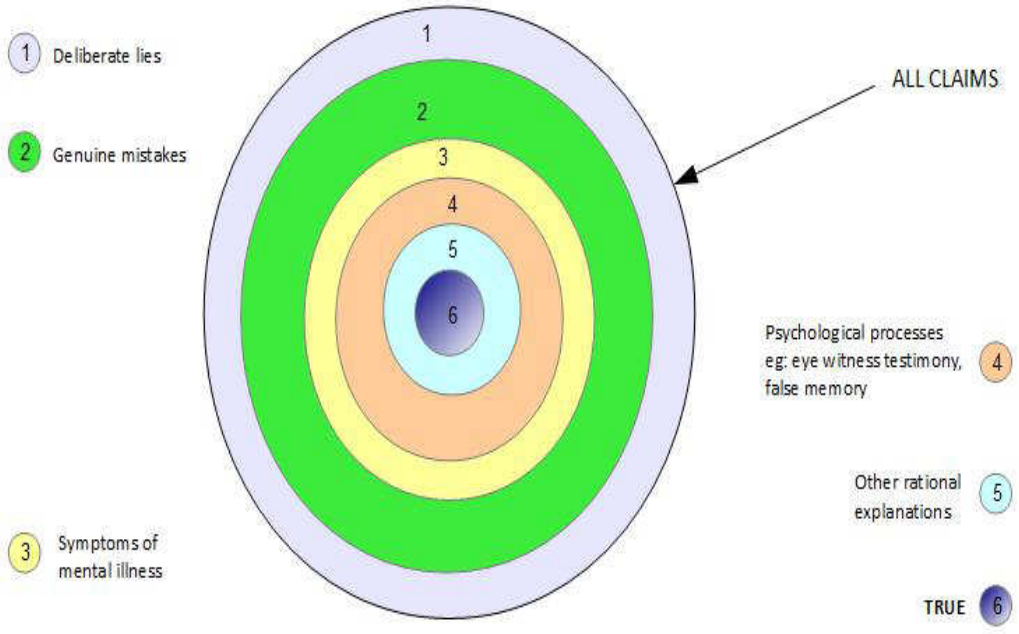


Figure 3.2 - Assessing paranormal claims.