

RESEARCH ASSISTANTS OR BUDDING SCIENTISTS? A REVIEW OF 96 UNDERGRADUATE STUDENT PROJECTS AT THE KOESTLER PARAPSYCHOLOGY UNIT

BY CAROLINE WATT¹

ABSTRACT. This paper reviews 96 undergraduate student projects supervised by Koestler Parapsychology Unit staff. Of these projects, 41.7% were unrelated either to psi or to paranormal experiences and beliefs; 18.8% were related to paranormal experiences and beliefs but contained no psi task. The paper focuses on the remaining 38 projects (39.6%) that included a psi task. Of these projects, 27 predicted overall significant performance on a psi task and 8 (29.6%) found the predicted results. Four small clusters of projects that used similar psi tasks were identified. The PK-RNG and EDA presentiment studies showed little consistency of psi task performance, and the majority (8 out of 9) obtained nonsignificant psi results. In contrast, significant results and relatively consistent effect sizes were found in 3 of the 5 ganzfeld studies, specifically those with participants who were “creative,” extravert, or who had practiced a mental discipline. None of the EDA-staring studies was statistically significant: however all 3 found effects in the predicted direction and of a reasonable magnitude ($r \approx 0.15$). These results are discussed in terms of what can be learned from student projects.

Particularly in Europe, academics with research interests in parapsychology are taking up posts within higher education institutions. Naturally these researchers become involved with teaching students and, as a result, many student projects are being conducted under their supervision. This raises the question of the evidential status of such projects. For instance, might they contain unreported flaws? Should they be included in meta-analyses of psi research? Furthermore, as there are far more student projects than there are parapsychologists, does the database of student projects form a significant file-drawer for parapsychology if these projects remain unpublished? To some extent these questions may be answered by taking into consideration the pedagogical context of the student projects: does the supervisor maintain strict control so that the student is essentially a research assistant, or is the student given some freedom to develop his or her scientific curiosity but also to make mistakes (and — we hope — learn from them)? This paper aims to encourage consideration of these important issues for modern-day parapsychology by looking in detail at the undergraduate student projects conducted at the Koestler Parapsychology Unit (KPU).

The KPU is a research group situated within the School of Philosophy, Psychology, and Language Sciences at the University of Edinburgh, Scotland. The KPU initially grew up around the Koestler Professor of Parapsychology, Robert Morris,² who held this position from 1985 until his death in 2004. The KPU consists of teaching staff, research staff, and postgraduate students whose area of interest and expertise is parapsychology and the psychology of anomalous experiences and beliefs. While Bob Morris was Koestler Professor, KPU researchers included (in alphabetical order) Deborah Delanoy, Peter Lamont, Julie Milton, Fiona Steinkamp, Paul Stevens, and Caroline Watt.

KPU members have always played a role in teaching psychology undergraduate and postgraduate students. The undergraduate psychology course at the University of Edinburgh is 4 years long. The single most substantial piece of work that undergraduate students produce is their final-year project and thesis. In this, students are able to apply their methodological and statistical training to topics in which they have a particular interest, under the supervision of staff with relevant expertise. Students can choose their thesis topic from a selection offered by staff. They are also free to approach staff with their own research idea, in which case the staff member may agree to supervise, so the degree to which the project topic is driven by a supervisor’s own research agenda varies depending on his or her pedagogical and research goals. Some may exert strict control so that the project addresses a tightly specified question in their research program and is methodologically sound. In this case, the student is essentially acting in the role of research assistant or coinvestigator. Others may prefer to encourage students to think creatively and follow their scientific curiosity. Here the project is less likely to contribute to a systematic research program, and

¹ This is a revised version of a paper presented by the author at the 2007 Parapsychological Association Convention. My grateful thanks go to Professors Richard Wiseman and Etzel Cardeña, and to my anonymous referees for their helpful comments on the paper.

² Prior to Morris’s arrival, Dr John Beloff for many years conducted parapsychological research at Edinburgh and supervised both undergraduate and PhD student projects. This paper does not survey the pre-Morris projects.

— in the hands of less able students — study quality may suffer. The project is written up as a thesis of around 11,000 words that is graded and contributes towards the student's overall degree grade.

To illustrate the issues raised by student parapsychology projects, this paper presents a descriptive overview of the final-year projects that have been supervised by KPU staff since its inception. These may act as a large number of pilot projects that might help guide future research. As these final-year projects are mostly unpublished, they may also place published studies in a wider context with respect to possible file-drawer effects. The KPU-supervised projects consist of two types: those that include a psi task and those that do not. As space is limited, I will only briefly describe the non-psi projects before going into the psi projects in greater detail.

METHOD

Study Retrieval

Every available 4th year thesis that had been supervised by KPU members was retrieved from the Koestler Library or directly from KPU members. Students are permitted to work in pairs (and, rarely, in groups of three or four) for data collection but have to write up their thesis independently, sometimes resulting in two or more theses written on the same group of data. In these cases, the thesis that was retrieved first was taken to represent the joint project.

Study Characteristics

Four main types of study characteristics were recorded.

The first is the main "demographic characteristics" of the thesis collection. The number of participants gives some indication of the statistical power of each study (which affects the potential of the study to detect a hypothesized genuine experimental effect to a statistically significant degree,) and the sex of participants gives an overall picture of the gender balance in these studies. Finding support for preplanned hypotheses indicates the "success" of the project (in terms of finding support for predicted effects and relationships), though of course a successful study in pedagogical terms could have null findings but be well designed, conducted, and reported.

The second concerns the research area of the project, including whether the study includes a psi measure, and a more detailed breakdown of the study topic.

The third provides a fuller picture of the characteristics of the psi studies in particular. One question of interest is whether the psi studies include a prediction of an overall psi effect (e.g., they include a "proof-oriented" hypothesis). In contrast, overall psi effects in process-oriented studies can be "washed out" if the experimental design includes conditions predicted not to be psi-conducive. As student projects often contain exploratory analyses, the extent to which the projects reported psi findings arising from post hoc analyses was also noted.

The fourth study characteristic of interest was whether the project has reached the public domain, either through journal publication or presentation at a Parapsychological Association conference (where details of method and analyses can be found in the conference Proceedings). In no cases was the full thesis published, but a version of the project could be written up for conference presentation or publication. If neither occurred, then the project is essentially in the "file-drawer." KPU student projects are also sometimes presented at the conference of the Society for Psychical Research. However, this conference provides only abstracts, so many methodological and statistical details are unavailable; one could argue that this lack of detail means the projects are essentially "unpublished."

RESULTS

Before I describe the results of this survey, some caveats are in order. First, the information provided here is meant to give a *descriptive*, not a *quantitative*, picture of the content of KPU student projects as described by the students themselves in their theses. This is not a meta-analysis because there are some inherent limitations to this database: the projects vary in detail, making it difficult to combine them. Second, although the projects were supervised by experienced members of the teaching and research staff, normally such supervision would not extend beyond advising the student on study design, helping provide facilities, and advising on analysis and write-up. It would therefore be rare for a supervisor to be present during actual data collection or to check that the students had correctly followed the planned procedure. In other words, in terms of the distinction I set up at the outset of this paper between directing research assistants and encouraging budding scientists, I would say that most of the KPU staff have preferred to follow the latter route. Third, time has passed, students and supervisors have departed, and in most cases we no longer have access to the studies' raw data. So we are required to take the study conclusions at face value and to trust that the students have been both truthful and proficient in how they have collected, analyzed, and

reported the data. Finally, although the supervisors may have given students feedback on the thesis contents during the writing process, the finalized theses have not been subjected to peer review.³

Demographic Characteristics of the Sample

A total of 127 students completed projects supervised by KPU members. Of these, 25 projects were jointly conducted by two or more students, giving 96 projects in all.

Year of project and supervisor. The projects span a 20-year period, from 1987 to 2007. Most were conducted in the 1990s and the vast majority (72/96 = 75%) were solely supervised by Robert (Bob) Morris (RM), with a further 12 (12.4%) being cosupervised by RM.

Participants. A total of 4,717 individuals participated in these projects. The mean number of participants per project was 49.9.⁴ The relatively large standard deviation for participant numbers (34.1) indicates a great deal of variability in project size; indeed the projects ranged in size from 12 to 228 participants (*Mdn* = 40; *Mode* = 40). This variability often reflected the methodology being used. For instance, the largest study was questionnaire-based and used factor analysis; therefore large numbers were needed for reliable results. Thirty-five (36.5%) of the projects gave no information about the sex of participants. Of the remainder, there were approximately equal numbers of male (*M* = 23.8, *SD* = 15.8) and female participants (*M* = 26.5, *SD* = 18.8). This suggests that the student experimenters were making some effort to go beyond the usual convenience sample of undergraduate psychology students. Otherwise we would expect to see a predominance of female participants, as females outnumber males by about 3:1 among the students' peer group.

Project "success." It turned out to be difficult to ascertain the success of each project in terms of finding support for preplanned hypotheses because most projects planned to test several hypotheses and did not clearly indicate a single main hypothesis. It was fairly common, therefore, for students to report mixed findings, with support for some hypotheses but not others. Of course, the more hypotheses tested, the greater the likelihood that some of these will achieve statistically significant support by chance alone.

Research Area

Psi or non-psi? The 96 projects were first categorized into two groups: those that included a psi task and those that did not. Just over one third of the projects (38/96 = 39.6%) included a psi task. Before looking more closely at these, the following section briefly describes the non-psi studies.

The non-psi studies. Fifty-eight projects (60.4%) did not include any psi task. These were further subdivided into topics that had a direct link with parapsychology and the psychology of paranormal experiences and beliefs, and those with no such link. It is perhaps interesting to note that, although this review concerns projects supervised by KPU members, 40 projects (41.7% of the total 96 projects) were not directly related to the paranormal. These projects covered research areas such as sports psychology, volition, and occupational psychology, largely reflecting Bob Morris's interest in performance enhancement. This was an area where Morris considered that parapsychology might have something to contribute to mainstream psychology, and vice versa. Example titles are:

Stress within organisations: An evaluation of a stress management course with the Scottish Police College.

An exploratory study into athletes' delay in attending two sports injury clinics.

The relationship between volition and self-esteem in alcoholism: A new model of volitional control.

There were 18 non-psi projects that *were* related to the paranormal (18.8% of the total 96 projects). These involved study of the psychology of paranormal beliefs and experiences, and factors affecting free-response ESP judging. Appendix A gives a full list of the year, titles, authors and supervisors of these projects. Example titles are:

The effect of varying levels of instruction/experience upon the judging of free-response ESP data.

Do people's beliefs in ESP affect the way they evaluate evidence confirming or disconfirming their beliefs?

The therapeutic consequences of visiting a medium.

³ A small number of projects have been written up for PA presentation, in which case the PA version has had some peer scrutiny. However, the current paper is based on the students' unreviewed theses.

⁴ This figure excludes one project that did not involve participants but consisted of a statistical analysis of over 800 NFL (National Football League) games.

The Psi Studies

The remainder of this paper concerns the 38 projects (39.6% of the total 96) that included a psi task. Appendix B gives a full list of the year, titles, authors, and supervisors of these projects. Almost three quarters of these projects included a hypothesis predicting overall significant performance on the psi task (27/38 = 71.1 %). The remaining 11 studies did not predict an overall effect, suggesting that these studies may have used contrasting conditions expected to “wash out” performance on the psi task when combined, such as psi-believers versus disbelievers (e.g., projects by West and Newell).

Overall predicted outcome. Of the 27 projects predicting overall significant performance on the psi task, 8 (29.6%) found a predicted significant effect. This is considerably greater than the 5% of studies expected to be significant at the .05 level by chance alone. Although some of these studies do report results that are only marginally significant, others report effects with much smaller *p*-values that are by definition even less likely to occur by chance alone, such as $p < .000001$ (Langbridge) and $p = .0005$ (Norfolk & Rose). However, most projects did not prespecify the analysis that would be used to evaluate the hypothesis, and some had more than one psi outcome (e.g., EDA presentiment and Blood Pulse Volume (BPV) presentiment in the case of Norfolk and Rose; conscious guessing and unconscious (physiological) responses in the case of Sah) and did not prespecify a single psi outcome measure. These factors are likely to inflate the likelihood of finding significant effects by chance alone.

Statistical power. Student studies tend to have relatively modest resources. This makes it difficult to have large numbers of participants, leading to reduced statistical power. This in turn makes it more difficult to obtain significant results. To estimate whether this might have been a factor for the 70% of projects that predicted but did not obtain overall significant psi task performance, I compared the studies' participant numbers. The studies that predicted and found overall significant scoring on the psi task had a mean of 34.9 ($SD = 10.2$) participants. Those that predicted and did not find significant scoring on the psi task had 36.8 ($SD = 11.3$) participants on average. This suggests that, compared to the significant studies, the nonsignificant studies are not being unduly hampered by low statistical power.

Post hoc psi findings. Over half of the psi projects (22/38 = 57.9%) reported significant post hoc findings related to psi task performance. This is perhaps not surprising because student projects often include a number of exploratory analyses and sometimes use novel methods that do not allow firm predictions to be made. We might expect that the students conducting the 19 studies that predicted but *did not find* overall significant psi task performance would be particularly tempted to data-mine in order to boost their study's findings. Sixty-three per cent of these projects (12/19) reported post hoc psi findings, compared to 50 % (4/8) of the studies that had predicted and found significant psi task performance. It appears, therefore, that there was a slight tendency for students to engage in more data-exploration when they did not find a predicted main effect.

Type of psi task. The projects used many and varied psi tasks. ESP/precognition tasks included dowsing for a hidden penny, presentiment of photographs of happy and angry faces, detecting the happy or neutral emotional state of a sender, as well as the better known ganzfeld free-response method and forced-choice studies with Zener cards. PK tasks included attempting to exert a psychokinetic influence over the fall of dice and over RNG-driven computer visual displays, including choosing which of four static rectangles would be selected by the computer, controlling a randomly moving spot's position relative to a circle, and influencing a balloon being driven by a fan onto spikes. DMILS studies included attempts to facilitate attention-focusing and performance on a choice reaction time task, and the autonomic detection of remote observation. In addition to the variety of psi tasks, most studies incorporated an investigation of various psychological characteristics, such as personality, creativity, volitional competence, volitional strategy, imagery strategy, belief/disbelief in ESP, meditation, and self-confidence. Given that we are dealing with only 38 psi studies with different combinations of the variables listed above (and more), the studies are not easily compared. Nevertheless there were some small clusters of studies that used fairly similar psi tasks, and these are described in the following sections in an attempt to see whether any informative patterns emerge. Each table presents the studies in chronological order; further details of date, title, and supervisor can be found in Appendix B.

Free-response ganzfeld ESP studies. There are five studies in this group, four of which have been presented at PA conventions (Morris, Cunningham, McAlpine, & Taylor, 1993; Symmons & Morris, 1997; Morris, Summers, & Yim, 2003). All used an automated ganzfeld procedure in which targets were automatically displayed in a distant room while the participant relaxed, experienced unpatterned sensory stimulation, and reported aloud his or her thoughts and impressions. The Symmons study used rhythmic auditory stimulation (drumming at 7 and 2 Hz) rather than white noise. Table 1 gives the principal characteristics of these studies.

Insert Table 1 around here

We can see from Table 1 that three of the five ganzfeldESP studies obtained above-chance hit-rates (mean hit-rate = 33.4%, mean effect size = 0.18⁵). Though the numbers are small and other covariates may account for the studies' outcomes, the nonsignificant studies used unselected student participants whereas the significant studies used participants selected for "creativity," extraversion, and prior psi experience or use of a mental discipline. Two of the studies had near-zero effects, whereas three had comparable effect sizes around .30 in magnitude, suggesting some consistency in the findings

Insert table 2 around here

RNG-PK studies. There were four studies that reported an overall PK outcome, as shown in Table 2. Participants were asked to attempt to influence a visual display (on a computer monitor) that was being driven by a random number generator (RNG). Three of the studies (Dumughn, Gentles, and Grice) used a PK task in which the participant presses the space bar to cause a dot to take a random step either remaining within or leaving a circle, with the dot always beginning in the centre of the circle. Each trial would terminate when the dot left the circle. In half of the trials, the participants were instructed to will the dot to stay within the circle (the IN condition) and in the other half they were to will the dot to leave the circle (the OUT condition). The number of steps taken was counted and significant psi scoring was noted if there were more steps for the IN than the OUT condition. Students reported the outcome of these studies by condition or, in the case of Lumley-Saville, per trial, rather than per bit. None of these studies obtained results approaching significance and effect sizes were rather diverse, suggesting inconsistent findings.

EDA-staring studies. There were three studies in which electrodermal activity (EDA) was monitored while participants were the subject of periods of remote observation (via camera) interspersed with periods of nonobservation. As shown in Table 3, none of the studies obtained significant results. Effect sizes were of comparable magnitude, though slightly lower when a convenience sample of undergraduate students was used.

Insert table 3 around here

EDA-presentiment studies. There were five EDA-presentiment studies, as shown in Table 4.⁶ In these studies, EDA was measured in the few seconds *prior to* exposure to emotive or neutral stimuli. The earliest of these studies (Norfolk & Rose) found a highly significant effect in the predicted direction; one found a nonsignificant effect in the predicted direction, and three found nonsignificant effects in the direction opposite to that predicted (i.e., lower skin conductance prior to exposure to emotive compared to control stimuli). Effect sizes were rather varied, suggesting inconsistent findings. Norfolk and Rose investigated and discounted the possibility that their results might be inflated by the use of anticipatory strategies by participants. However, there is another reason to doubt the reliability of Norfolk and Rose's findings, as explained further in the discussion section.

Insert table 4 around here

Process-Oriented Studies

As mentioned above, 11 studies did not predict an overall significant psi effect. Most of these were process-oriented studies that either correlated individual difference variables with psi outcome or compared two conditions, one of which was expected to be more psi-conducive than the other. Again, this was a heterogeneous group in terms of the research question; however two small clusters of similar studies can be identified. Four looked at sheep-goat effects (one predicting a reversed SGE, replicating Lovitts, 1981), and three found the predicted significant

⁵ Formulae for effect size conversion can be found in Rosenthal and Rosnow. (1991).

⁶ The study by Norfolk & Rose had two psi outcomes: EDA presentiment and BPV presentiment. To avoid reporting two psi outcomes for a single study, I report only the EDA results here because this was the first hypothesised psi outcome.

relationship between belief and ESP (West, Newell, and Gentles). Four looked at aspects of volition and psychokinesis task performance, and two found significant relationships (MacLulich and Prince).

File-Drawer Issues

Of the 96 projects, only one has led to a journal publication (Watt, Wilson, & Watson, 2007). Not surprisingly, none of the nonpsi projects that are unrelated to the paranormal has been presented at PA conventions. More surprisingly, none of the nonpsi projects that *are* related to the paranormal has been presented at PA conventions. Seven psi studies (18.4%) have been presented at PA conventions^{7,8} (with two being presented in a single paper). These are listed in Table 5; full references can be found in the reference list. It appears that there is a tendency to select for PA presentation studies that have found significant results for the psi task, because the proportion of significant studies presented at the PA (5/7 = 71.4%) is considerably larger than the overall proportion of studies that predicted and found significant overall psi effects (8/27 = 29.6%).

Insert table 5 around here

DISCUSSION

This paper reviewed 96 undergraduate student projects supervised by members of the Koestler Parapsychology Unit at Edinburgh University and conducted between 1987 and 2007. Of these projects, 41.7% were unrelated either to psi or to paranormal experiences and beliefs. Around 18.8% were related to paranormal experiences and beliefs but contained no psi task. Of the 38 projects (39.6%) that included a psi task, 27 predicted overall significant performance on a psi task, of which 8 (29.6%) found the predicted significant overall positive psi scoring. However, some of these projects had more than one psi task and did not specify a single outcome measure. Interestingly, there was little difference between the number of participants (and hence statistical power) for the “successful” (in terms of finding a predicted overall psi effect) and the “unsuccessful” studies. However, with only around 35 participants each, the psi studies would conventionally be regarded as low-powered. This is not particularly surprising, given that these were student projects with necessarily limited resources.

Although the projects were very diverse (see Appendix B), in an attempt to discern informative trends, four small clusters of projects were identified that had used similar psi tasks: PK-RNG, EDA presentiment, ganzfeld ESP, and EDA staring studies. The first two tended not to show any consistency of performance on the psi task, with the majority (8/9 studies) obtaining nonsignificant psi results. Perhaps this is because all the PK-RNG studies analysed their data on a per-condition or per-trial basis, rather than on a per-bit basis. Therefore the power of the student studies, in terms of number of trials, is very low compared to the PK-RNG studies meta-analysed by Bösch, Steinkamp, and Boller (2006) that had thousands and sometimes millions of bits, and that found a very small but significant overall effect size. However, effects for the PK-RNG and presentiment studies often occurred in the direction opposite to that hypothesized, suggesting that inconsistent results might continue to be found even with improved statistical power. Of course the alternative possibility is that these studies simply did not support the hypothesis of the existence of psi because, even if psi exists, it is not readily detected using PK-RNG or EDA presentiment methods.

As mentioned above, a cause for concern arises over the one significant EDA-presentiment study (Norfolk & Rose). The study by Dimmock and Tawse, conducted 1 year later as a replication of Norfolk and Rose, initially found highly significant results ($z = 5.57$). However, on further investigation Dimmock and Tawse found that this was due to a faulty data analysis program; when the data were manually analyzed, chance results were obtained. Based on the contents of the students’ theses, we cannot rule out the possibility that, although Norfolk and Rose explored and ruled out anticipatory strategies as causing their sizeable apparent presentiment effect ($z = 3.28$, effect size = 0.49), there might have remained a problem with their data analysis program as it is quite likely that the same or a similar program was used in both studies.

⁷ The studies by Cunningham and McAlpine were also reported at the PA by Radin, McAlpine, & Cunningham (1993) in the context of geomagnetic effects. The original projects had not planned to investigate geomagnetic effects, so I am taking the Morris et al. (1993) paper as representing these projects.

⁸ For the record, I am aware of one student project that falls outside the consideration of this paper because it was a third- year undergraduate project on staring detection and it was not supervised by KPU staff. It was presented at a PA convention as a research brief (Lee et al., 2005; see reference list for details).

More reliable findings came from the ganzfeld and remote staring studies. Significant results and relatively consistent effect sizes were found in the majority of the ganzfeld studies, specifically those with selected participant populations, such as persons who were “creative,” extraverted, or had practiced a mental discipline. This trend, though based on only a small sample of studies, is consistent with the high scoring found in previous (Schlitz & Honorton, 1992) and subsequent (Dalton, 1997) ganzfeld studies with creative participants, and with Milton and Wiseman’s (1999) meta-analysis that found above-chance scoring for participants who practiced mental disciplines. It seems that the use of selected participants is preferable for experimenters conducting proof-oriented ganzfeld ESP research. Furthermore, a small number of the process-oriented projects found evidence in support of the sheep-goat effect. This reminds us of one of parapsychology’s more reliable findings (e.g., Lawrence, 1993) and suggests that if parapsychologists want to maximize the likelihood of obtaining significant psi results then they should also select participants on the basis of their belief in the paranormal. It is perhaps worth emphasizing that even if researchers believe that the evidential basis for psi has already been well established (and therefore that there is no point in doing further proof-oriented research), we can nevertheless learn more effectively *how* psi works by maximising the psi-conduciveness of the experimental setting.

There were only three EDA-staring studies. None was statistically significant; however two were quite low-powered with only 28 participants each, making statistical significance difficult to achieve. All three found effects in the predicted direction and of a magnitude ($r \approx 0.15$) larger than that found in the Schmidt, Schneider, Utts, and Walach (2004) meta-analysis of 15 remote staring studies (with the latter finding a mean effect size $d = 0.13$, which converts to $r = 0.07^9$). However, the two databases are not independent because Schmidt et al. retrieved unpublished studies, including two of the three studies reported here (Howat, and Juniper & Edlmann). Nevertheless, particularly given that remote staring detection is a commonly reported real-life psi experience (Sheldrake, 2005) I would echo Schmidt et al.’s recommendation that further high quality research be conducted using this method.

This survey also suggests that there is a tendency to select for PA presentation projects that have obtained evidence supporting the psi hypothesis. This is not surprising because positive results are more interesting to present and to hear than null results, just as they are with journal publications. But this tendency does give an unrepresentative picture of the degree of support these projects have found for the psi hypothesis. Nonetheless, the dataset as a whole found a sizeable proportion of significant results, and thus it could be argued that it is overall supportive of a psi hypothesis.

This paper has two goals. The first was to put on record — to bring out of the file-drawer — the principal characteristics and findings of KPU student projects. I turn now to the broader goal, which is to consider the evidential status of mostly unpublished student projects at the KPU and other educational institutions. Given the already-mentioned limitations of taking the “budding scientists” pedagogical route, such as the actual data collection and analysis being largely the responsibility of the students alone, some might feel that these projects’ findings are less reliable than studies conducted by more experienced researchers or by students essentially acting as research assistants. Meta-analysts’ practices differ because there is a trade-off between avoiding file-drawer effects, and including possibly unreliable studies. Schmidt et al. (2004) included some of the unpublished studies reported here but also coded the studies on a number of different quality indices, thus allowing the effects of low quality studies to be quantified to some extent. Milton & Wiseman (1999) included the ganzfeld studies by McAlpine and Cunningham, not in their original unpublished form but as they were presented in the PA Proceedings. One would expect the reliability of published papers to be higher because they have undergone peer review. If students have been given a relatively free rein and their projects are unpublished, I would not recommend using such projects in a meta-analysis unless they are accompanied by thorough quality-coding. Even then, the project write-ups may conceal errors in procedure and analysis that could remain undetected by quality coding. Of course the same can apply to even the most experienced researcher, but we may assume that the trainee researcher is more error-prone.

Aside from the question of poor quality methodology, some might suspect that students would be motivated to simulate positive results because these findings are easier to report and are likely to be viewed more favorably by markers. This is an unfalsifiable hypothesis, so it is not very useful. However students using automated data collection would find it relatively difficult to simulate positive results, and the fact that only about 30% of the KPU-supervised psi studies predicted and then reported finding overall psi effects suggests that cheating is not widespread among these students. Furthermore, conscious or nonconscious biases during data collection and analysis may also be present among researchers who may have equal or greater stakes in producing significant results than students.

It has been difficult to find overall trends in this review given the miscellaneous nature of the bulk of the student projects. Such is almost a microcosm of parapsychology as a whole. The great strength of student projects is that they can be very innovative as the students themselves can come up with fresh research ideas. Symmons’ ganzfeld project using different drumming frequencies rather than white noise as auditory stimulation is one example

⁹ Using the formula $r = d/\sqrt{(d^2+4)}$

of a successful variation upon an established methodology. So we could regard such projects as risk-taking pilot studies that might suggest new lines of research for subsequent systematic investigation, or that explore modifications of well-known methodologies. However, if we want these projects to contribute to the database of formal psi studies, they need to be both closely supervised and systematic so that they may be published within a wider research context. At the same time, this will bring them out of the file-drawer. A few of the KPU student projects, particularly the ganzfeld studies with selected participants, are relatively programmatic. We could go further to encourage systematic student research, but that is likely to be at the cost of flexibility and novelty, and would not allow the student to follow his or her own initiative.¹⁰ In her excellent article on how educational activities in parapsychology can be used to recruit new researchers, veteran instructor Gertrude Schmeidler (1999) describes how she advises her students to conduct replication studies. But, she says, “the best part of that advice is that the bright students don’t take it” (p.29). It is down to the educator him or herself to make a judgment as to whether pedagogical goals can be satisfied simultaneously with wider research goals and, if not, which should take priority.

REFERENCES

- BÖSCH, H., STEINKAMP, F., & BOLLER, E. (2006). Examining psychokinesis: The interaction of human intention with random number generators — a meta-analysis. *Psychological Bulletin*, **132**, 497–523.
- BRADY, C., & MORRIS, R. L. (1997). Attention focusing facilitated through remote mental interaction: A replication and exploration of parameters. *Proceedings of Presented Papers: The Parapsychological Association 40th Annual Convention*, 73–91.
- DALTON, K. (1997). Exploring the links: Creativity and psi in the ganzfeld. *Proceedings of Presented Papers: The Parapsychological Association 40th Annual Convention*, 119–134.
- DELANOY, D. L., & SAH, S. (1994). Cognitive and physiological psi responses to remote positive and neutral emotional states. *Proceedings of Presented Papers: The Parapsychological Association 37th Annual Convention*, 128–138.
- LAWRENCE, T. R. (1993). Gathering in the sheep and goats: A meta-analysis of forced-choice sheep-goat ESP studies, 1947–1993. *Proceedings of Presented Papers: The Parapsychological Association 36th Annual Convention*, 75–86.
- LEE, B., HADJIOSIF, M., WEIDLICH, D., BARLOW, L., CRICHTON, C., MOOREHEAD, C., ET AL. (2005). Personality and perception in staring detection and experience. *Proceedings of Presented Papers: The Parapsychological Association 45th Annual Convention*, 407–409.
- LOVITTS, B. E. (1981). The sheep-goat effect turned upside down. *Journal of Parapsychology*, **39**, 289–296.
- MILTON, J., & WISEMAN, R. (1999). Does psi exist? Lack of replication of an anomalous process of information transfer. *Psychological Bulletin*, **125**, 387–391.
- MORRIS, R.L., CUNNINGHAM, S., MCALPINE, S., & TAYLOR, R. (1993). Toward replication and extension of autoganzfeld results. *Proceedings of Presented Papers: The Parapsychological Association 36th Annual Convention*, 177–191.
- MORRIS, R. L., SUMMERS, J., & YIM, S. (2003). Evidence of anomalous information transfer with a creative population in ganzfeld stimulation. *Proceedings of Presented Papers: The Parapsychological Association 46th Annual Convention*, 116–131.
- RADIN, D. I., MCALPINE, S., & CUNNINGHAM, S. (1993). Geomagnetism and psi in the ganzfeld. *Proceedings of Presented Papers: The Parapsychological Association 36th Annual Convention*, 132–142.
- ROSENTHAL, R., & ROSNOW, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.) New York: McGraw-Hill.
- SCHLITZ, M. J., & HONORTON, C. (1992). Ganzfeld psi performance within an artistically gifted population. *Journal of the American Society for Psychical Research*, **86**, 93–98.
- SCHMEIDLER, G. R. (1999). Recruiting for research. In *Education in Parapsychology* (pp. 17-33). New York: Parapsychology Foundation. (Edited from the original article of the same name, *Parapsychology Review*, 1976, **7**(1), 1-5).
- SCHMIDT, S., SCHNEIDER, R., UTTS, J., & WALACH, H. (2004). Distant intentionality and the feeling of being stared at: Two meta-analyses. *British Journal of Psychology*, **95**, 235–247.

¹⁰ A similar debate could probably be held about the research of PhD students, where supervisors vary widely in the extent to which they exert tight control over the student’s research topic and experimentation. The PhD dissertation is another class of student research that may contribute to parapsychology’s research database but, if unpublished, also add to its file-drawer.

- SHELDRAKE, R. (2005). The sense of being stared at. Part 1: Is it real or illusory? *Journal of Consciousness Studies*, **6**, 10–31.
- SYMMONS, C., & MORRIS, R. L. (1997). Drumming at seven Hz and automated ganzfeld performance. *Proceedings of Presented Papers: The Parapsychological Association 40th Annual Convention*, 441–453.
- WATT, C., FRASER, C., & HOPKINSON, A. (2006). “Psychic DMILS”: Attempted remote facilitation of performance in an ESP game, and an exploration of gender differences. *Proceedings of Presented Papers: The Parapsychological Association 49th Annual Convention*, 226–236.
- WATT, C., WILSON, S., & WATSON, L. (2007). Cognitive and psychological mediators of anxiety: Evidence from a study of paranormal belief and perceived childhood control. *Personality and Individual Differences*, **42**, 335–343.

*Koestler Parapsychology Unit
School of Philosophy, Psychology and Language Sciences
University of Edinburgh, UK
caroline.watt@ed.ac.uk*

APPENDIX A

PROJECTS RELATED TO THE PARANORMAL THAT DID NOT INCLUDE A PSI TASK

Year	Title	Author(s)	Supervisor(s)
1987	An experimental evaluation of a belief in psi	Reed	Morris
1987	Self disclosure through non-verbal leakage — How to read minds through what bodies are doing	Lindsay	Morris
1989	Attitude and assessment: The role of personal belief in objective evaluation	Roe	Morris & Milton
1989	The effect of varying levels of instruction/experience upon the judging of free-response ESP data	McKie	Delanoy
1993	Perceived control, probability judgement and belief in psi phenomena	McClure	Morris
1994	Factors affecting the attribution of meaning to coincidence in judging free-response ESP tasks	Austin	Morris
1994	The attaching of significance to coincidence	Cupples	Morris
1995	The measurement of perceptual defence: Further exploration using Pandora's Box	Robins	Morris & Watt
1995	The coincidence prone personality and the attribution of meaning to coincidence	Cheeseman	Morris
1995	Extremity of belief: rigidity, resistance to shift, and revising attributions made to anecdotal coincidence	Levy	Morris
2000	The effects of a lucky event and increased illusion of control on people's belief in luck	Wright, Wright & Adams	Watt
2001	Do people's beliefs in ESP affect the way they evaluate evidence confirming or disconfirming their beliefs?	Cook	Morris
2003	An exploration of the relationship between perception of control in childhood, locus of control and paranormal belief	Watson & Wilson	Watt
2004	Evaluations of television mediums: The effects of prior belief and presentation style (one-sided versus two-sided)	Flint & Warmington	Watt
2004	The effect of credibility variation in television disclaimers, with reference to television mediums	Anderson & Nesbitt	Watt
2005	Paranormal belief: Locus of control, defence mechanisms, and coping styles	Campbell	Watt
2007	Developing a new measurement of paranormal belief using the Implicit Association Test	Danelian & Stewart	Watt
2007	The therapeutic consequences of visiting a medium	O'Neill	Watt

APPENDIX B

PROJECTS THAT INCLUDE A PSI TASK

Year	Title	Author(s)	Supervisor(s)
1989	A first revision of Lovitts' (1981) sheep-goat effect turned upside down	West	Morris
1989	Inferring correspondence in a free-response ESP judging task: Effects of training and judging strategy	Fairfoul	Delanoy & Morris
1990	A novel self-report measure of volitional styles and performance on a computer-based psychokinesis task	MacLulich	Morris
1992	The attribution of meaning to coincidence	Newell	Morris
1993	ESP, personality and belief: An interactional approach	Mekie	Morris & Delanoy
1993	The effects of creative ability and individual differences on extrasensory perception in the ganzfeld	Cunningham	Morris
1993	An experimental investigation into the effects of attitude and target type on psi in the ganzfeld: An attempt at replication and extension	McAlpine	Morris
1993	Applying volitional imagery strategies learned in a psychokinesis game to managing the smoking habit	Dumughn	Morris
1993	Imagery, volition, and belief factors in relation to scoring on a PK computer game	Gentles	Morris
1993	Investigating the effects of two different volitional strategies, and level of competitive anxiety, on performance in a computer-based test of psychokinesis	Grice	Morris
1993	Measuring ESP using dowsing with comparison to personality traits	Hayes	Morris
1993	The effect of arousal on an exploratory extrasensory perception experiment	Boswell	Delanoy
1994	Remote staring detection and personality correlates	Howat	Morris & Delanoy
1994	The extra-sensory perception of emotions as measured by	Sah	Delanoy & Morris

	cognitive (conscious) and physiological (unconscious) responses		
1996	An experimental parapsychology study into the concept of interconnectedness	Lort-Phillips	Morris
1996	Volitional competence and performance on a PK task	Sullivan	Morris
1997	An examination of the effects of an auditory stimulus of drumming at the theta frequency on an automated Ganzfeld experiment	Symmons	Morris
1997	An investigation into the phenomenological dimensions of the meditative state and their associations with ESP	Baber	Morris
1997	Attention focusing facilitated through remote mental interaction: A replication and exploration of parameters	Brady	Morris
1997	Dream ESP: A comparison of lucid and normal dreaming using complex and simple targets	Langbridge	Morris
1998	EMG biofeedback in facial muscles expected to enhance detection of future affective states in those muscles	Kettle	Morris
1998	Asymmetry of facial EMG in posed and spontaneous facial expressions of emotion: A psychophysiological study with reference to presentiment effects	McGee & Menzies	Morris
1998	The relationship between perceptual vigilance/defensiveness and psychophysiological responses to remote staring whilst in the ganzeld condition	Juniper & Edlmann	Morris
1999	An investigation into the effect of a sender in a card guessing experiment	Moss	Steinkamp
1999	Performance on a choice reaction time task facilitated through remote mental influence	Roberts & Moore	Morris, Delanoy & Watt
1999	Can future emotions be perceived unconsciously? An investigation into the presentiment effect with reference to extraversion	Norfolk & Rose	Morris & Delanoy
1999	The effect of meditation on PK performance in novice meditators	Lumley-Savile	Morris

2000	Anomalous arousal levels in relation to sensation seeking: Presentiment or anticipatory strategies?	Dimmock & Tawse	Morris & Stevens
2001	Imagery in the ganzfeld: The effects of extroversion and judging instructions upon imagery characteristics	Colyer	Morris
2001	How does self-confidence relate to PK performance on a dice-influencing task	Prince	Morris
2003	Unconscious precognition, experimenter effects, and participant individual differences: An EDA study	Crichton, Lowe, Dowding, Wilson & Wright	Morris & Stevens
2003	Investigating unconscious telepathic communication of a remote auditory startle and the effect of auditory entrainment	Munckton & Westcott	Morris
2003	Evidence of anomalous information transfer with a creative population in the ganzfeld	Summer & Yim	Morris
2004	Differences in baseline skin conductance to neutral and arousing stimuli with respect to normal and applied magnetic fields	Costley & Gregory	Morris & Stevens
2004	The effects of different auditory stimuli and of relaxation in evoking presentiment	Cochrane & Wilson	Morris
2004	Exploring the why of creative psi in musical agent receiver pairings: Does unconscious psi perception mediate psi hitting?	Cohen & Reid	Morris
2005	Psychic DMILS: Can remote helping facilitate performance on an ESP game?	Hopkinson & Fraser	Watt
2005	Facilitation of attention focusing by a remote helper: comparison of informed and uninformed conditions	Martin & Miller	Watt

TABLE 1
CHARACTERISTICS OF FREE-RESPONSE GANZFELD-ESP STUDIES CONDUCTED AS STUDENT PROJECTS^a

Author(s)	Participant type	Hits/Trials (MCE = 25%)	z(one-tailed p)	Effect size z/√N
Cunningham	Creative & extravert undergrad “sheep”	13/32 = 40.6%	1.84 p = .03	.33
McAlpine	Unselected undergrads	8/32 = 25%	0	0
Symmons	Mostly undergrads selected for prior psi experience & mental discipline	21/51 = 41.2%	2.51 p = .006	.35
Colyer	Unselected undergrads	9/40 = 22.5%	-0.18	-.03
Summers & Yim	Selected artists & musicians	15/40 = 37.5%	1.643 p = .05	.26

Note When a negative value is shown, this indicates the results are in the direction opposite to that predicted; one-tailed *p*-values are only given when the results are in the predicted direction.

^a All projects except for McAlpine used dynamic video clips as targets. McAlpine used static and dynamic video clips.

TABLE 2
CHARACTERISTICS OF PK-RNG STUDIES CONDUCTED AS STUDENT PROJECTS THAT REPORTED AN OVERALL PK OUTCOME

Author(s)	Participant type	RNG type	PK outcome	Effect size
Dumughn	Undergraduate female smokers	Algorithm	<i>t</i> (24) = -1.601	-0.31 ^a
Gentles	Unselected undergrads	Algorithm	<i>t</i> (59)=1.21 p=0.12	0.16 ^a
Grice	Undergrads selected for sporting experience	Algorithm	<i>t</i> (39)=-0.05	-0.008 ^a
Lumley-Saville	Undergrads selected for no meditation experience	Unreported	Mean hits = 40.015 (80 trials, MCE = 40, <i>z</i> = .003)	.0003 ^b

Note. When a negative value is shown, this indicates the results are in the direction opposite to that predicted; one-tailed *t*-values are only given when the results are in the predicted direction.

^a $r = \sqrt{t^2/t^2 + df}$

^b (*z*/√*N*)

TABLE 3
CHARACTERISTICS OF EDA-STARING STUDIES CONDUCTED AS STUDENT PROJECTS

Author(s)	Participant type	Identity of starrer	Psi outcome	Effect size
-----------	------------------	---------------------	-------------	-------------

				$r = \sqrt{(t^2/t^2+df)}$
Howat	General volunteers	Experimenter	$t(27) = 0.905^a$, $p = .38$, two-tailed	0.17
Lort-Phillips	Involved with "dance subculture"	Experimenter	$t(27) = 0.96^b$, $p = .17$, one-tailed	0.18
Juniper & Edlmann	Unselected undergraduates	Experimenters	$t(65) = 0.917$, $p = .36$, two-tailed	0.11

^aNo directional prediction was made in this study; EDA was more activated during stare than no-stare periods.

^bAs predicted, EDA was calmed during stare compared to no-stare periods.

TABLE 4

CHARACTERISTICS OF EDA-PRESENTMENT STUDIES CONDUCTED AS STUDENT PROJECTS

Author(s)	Participant type	Stimulus type	Psi outcome	Effect size
Norfolk & Rose	Unselected undergraduates ($N = 43$)	Extreme (positive & negative) versus calm pictures	Wilcoxon $z = 3.28$, $p = .0005$,	0.49
Dimmock & Tawse	Unselected undergraduates ($N = 36$)	Extreme (positive & negative) versus calm pictures	Wilcoxon $z = -1.6$	-0.24 ^a
Crichton, Lowe, Dowding, Wilson & Wright	Unselected undergraduates	Acoustic startle versus silent control	$t(71) = -1.536$,	-0.18 ^b
Costley & Gregory	Unselected undergraduates	Emotionally arousing versus neutral pictures	$t(36) = -1.188$	-0.04 ^b
Cochrane & Wilson	Undergrads selected for no meditation experience	Arousing versus neutral sounds	$F(1,36) = 0.555$, $p = .461$	0.12 ^c

Note. When a negative value is shown, this indicates the results are in the direction opposite to that predicted; one-tailed p -values are only given when the results are in the predicted direction.

^a (z/\sqrt{N})

^b $r = \sqrt{(t^2/t^2+df)}$

^c $\sqrt{[F(dfbet)/F(dfbet)+dfwithin]}$

TABLE 5

PROJECTS INCLUDING A PSI TASK THAT HAVE BEEN PRESENTED AT PA CONVENTIONS

Project Author(s)	PA Proceedings citation	Overall significant psi task performance found?
Cunningham	Morris, Cunningham, McAlpine & Taylor (1993)	Yes
McAlpine	Morris, Cunningham, McAlpine & Taylor (1993)	No
Sah	Delanoy & Sah (1994)	Yes

Symmons	Symmons & Morris (1997)	Yes
Brady	Brady & Morris (1997)	Yes
Summers & Yim	Morris, Summers, & Yim (2003)	Yes
Hopkinson & Fraser	Watt, Fraser, & Hopkinson (2006)	No