Eastern Illinois University
The Keep

Masters Theses | Student Theses & Publications

2018

An Analysis of Extended Strings of Hits and Misses in a Mental Telepathy Task

Gregory R. Peck
Eastern Illinois University

This research is a product of the graduate program in Clinical Psychology at Eastern Illinois University. Find out more about the program.

Recommended Citation
Peck, Gregory R., "An Analysis of Extended Strings of Hits and Misses in a Mental Telepathy Task" (2018). *Masters Theses*. 3445.
https://thekeep.eiu.edu/theses/3445

This is brought to you for free and open access by the Student Theses & Publications at The Keep. It has been accepted for inclusion in Masters Theses by an authorized administrator of The Keep. For more information, please contact tabruns@eiu.edu.

The Graduate School
EASTERN ILLINOIS UNIVERSITY

Thesis Maintenance and Reproduction Certificate

FOR: Graduate Candidates Completing Theses in Partial Fulfillment of the Degree
Graduate Faculty Advisors Directing the Theses

RE: Preservation, Reproduction, and Distribution of Thesis Research

Preserving, reproducing, and distributing thesis research is an important part of Booth Library's responsibility to provide access to scholarship. In order to further this goal, Booth Library makes all graduate theses completed as part of a degree program at Eastern Illinois University available for personal study, research, and other not-for-profit educational purposes. Under 17 U.S.C. § 108, the library may reproduce and distribute a copy without infringing on copyright; however, professional courtesy dictates that permission be requested from the author before doing so.

Your signatures affirm the following:

- The graduate candidate is the author of this thesis.

- The graduate candidate retains the copyright and intellectual property rights associated with the original research, creative activity, and intellectual or artistic content of the thesis.

- The graduate candidate certifies her/his compliance with federal copyright law (Title 17 of the U. S. Code) and her/his right to authorize reproduction and distribution of all copyrighted materials included in this thesis.

- The graduate candidate in consultation with the faculty advisor grants Booth Library the nonexclusive, perpetual right to make copies of the thesis freely and publicly available without restriction, by means of any current or successive technology, including but not limited to photocopying, microfilm, digitization, or internet.

- The graduate candidate acknowledges that by depositing her/his thesis with Booth Library, her/his work is available for viewing by the public and may be borrowed through the library's circulation and interlibrary loan departments, or accessed electronically. The graduate candidate acknowledges this policy by indicating in the following manner:

✓ Yes, I wish to make accessible this thesis for viewing by the public

______ No, I wish to quarantine the thesis temporarily and have included the ***Thesis Withholding Request Form***

- The graduate candidate waives the confidentiality provisions of the Family Educational Rights and Privacy Act (FERPA) (20 U. S. C. § 1232g; 34 CFR Part 99) with respect to the contents of the thesis and with respect to information concerning authorship of the thesis, including name and status as a student at Eastern Illinois University. I have conferred with my graduate faculty advisor. My signature below indicates that I have read and agree with the above statements, and hereby give my permission to allow Booth Library to reproduce and
gnature indicates concurrence to re

Graduate Candidate Signature

Gregory Peck
Printed Name

Clinical Psychology
Graduate Degree Program

Faculty Adviser Signature

Russell Gruber
Printed Name

4-13-2018
Date

Please submit in duplicate.

An Analysis of Extended Strings of Hits and Misses in a Mental Telepathy Task

(TITLE)

BY

Gregory R. Peck

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

IN THE GRADUATE SCHOOL, EASTERN ILLINOIS UNIVERSITY
CHARLESTON, ILLINOIS

Spring 2018

YEAR

I HEREBY RECOMMEND THAT THIS THESIS BE ACCEPTED AS FULFILLING
THIS PART OF THE GRADUATE DEGREE CITED ABOVE

THESIS COMMITTEE CHAIR DATE

THESIS COMMITTEE MEMBER DATE

THESIS COMMITTEE MEMBER DATE

THESIS COMMITTEE MEMBER DATE

THESIS COMMITTEE MEMBER DATE

An Analysis of Extended Strings of Hits and Misses in a Mental Telepathy Task

Clinical Psychology Master's Thesis 2018

Gregory R. Peck

Eastern Illinois University

Table of Contents

Acknowledgements

I'd like to thank Dr. Russell Gruber for offering me the opportunity to research mental telepathy using his lab data. Researching what most would consider to be esoteric topics was a big part of my decision to enroll in the graduate program at Eastern Illinois University, and I am thankful to have been involved in this project. I'd also like to thank my committee members, Dr. Bernas and Dr. Addison, for their input on how the research findings in this thesis could best be analyzed and presented. A heartfelt thanks goes out to my family and to the 2015 Clinical Psychology cohort, your support and encouragement has kept me motivated to stay in the program, even when the going got tough and many people would have quit. Last but not least, I'd like to thank my dear friend Luna for always being there for me throughout my time as a college student; because of your tireless support, I have achieved things that I once thought I could only dream of.

Abstract

The current study was designed to reanalyze archival data from the Eastern Illinois University parapsychology lab to determine the frequency of strings of hits and misses. Four datasets were included in the analyses, comprising a total of 348 participants. It was hypothesized that there would be more strings of hits and/or misses than expected by chance; there would be a difference between the number of strings of hits and the number of strings of misses; and participant scores on a 7-point belief in psi scale would be related to the frequency of hits and misses. Results indicated that no significant variance from chance was found for strings of hits or strings of misses, nor was there a significant difference in the overall number of strings. However, when considering the effect of belief, high belief participants scored significantly more strings of hits while low belief participants scored significantly more strings of misses (strings of two: $p = .0016$; strings of three: $p = .00014$). The correlation between belief and psi task performance also increased with longer strings, but tapered off after five consecutive hits or misses. The results of the study suggest that categorizing participants by belief may be an important factor in obtaining significant experimental findings. Additionally, with regard to belief in luck, it is speculated that gamblers may in fact accurately perceive strings of good luck and bad luck, which may promote the desire to gamble.

An Analysis of Extended Strings of Hits and Misses in a Mental Telepathy Task

Psi is the general term for a group of mental phenomena that have no established scientific explanation. It includes clairvoyance ("clear-seeing", or perceiving that which a person has no way of perceiving with their normal five senses) and telepathy (communication between minds). One way to think about psi effects is to examine them in terms of conditional probability, and indeed, the analysis I have conducted relies on some of the basic assumptions of probability. For example, if a particular outcome is expected to occur 25% of the time, we would expect that outcome to occur one in every four times that the contingent action was taken. However, probability theory tells us that the one in four chance is applied to *each individual attempt* rather than a group of four attempts (or however many we are considering). This means that, potentially, one can make hundreds or even thousands of attempts and not obtain the desired outcome. This does not contradict probability, however, because the probability of obtaining the desired outcome does not change. In a parapsychology experiment, we might see patterns of hits and misses closer to those obtained by recording the outcome of 100 coin flips. While we would expect a roughly equal number of hits and misses by chance, we might end up with a number that is skewed toward one side or another, and we might see that the coin landed on tails many times in a row before it landed on heads (and perhaps vice-versa). If the coin landed on tails 25 times in a row, most people would consider such an outcome highly improbable, as we tell ourselves that such a high number of tails in a row should not happen if the coin is fair. None of these improbable occurrences actually contradict the rules of probability, however, because the probability does not change across attempts. These improbably long strings of hits and misses are the crux of my analysis, because in the case of parapsychology experiments, a large number of hits or misses in a row may be representative of psi effects and not simply the result of a chance outcome. One theory about psi holds that it may be a manifestation of social bonds between humans (Sheldrake, 2014; Yarnall, 2003; Conner, 2009). Although psi is often elusive in nature, which has vexed parapsychology researchers for decades, its effects are often statistically significant and they have been demonstrated in parapsychology labs around the globe. Some of the research that has been done in the past 60 years is presented in this thesis, with the hope that it will give the reader a clearer picture of how psi manifests itself and how it is commonly detected and measured in experimental settings. Some of this research has been carried out at Eastern Illinois University (EIU) by Gruber, and I will also be presenting and discussing some of the parapsychology theses that have been written about his lab data since 2000. This thesis is primarily concerned with analyzing strings of "hits" and "misses" made by participants during a series of telepathy experiments conducted in EIU's parapsychology lab. The basic lab methodology will be presented and discussed, followed by a description of the analyses that were carried out along with the statistical results that were obtained.

Belief in Psi

Psi effects are often capricious in nature, that is, they seem to appear without following any established logical rules or patterns. However, the literature suggests that those who believe in psi tend to manifest psi effects more often.

Schmeidler and Murphy (1945) hypothesized that those who believe in extrasensory perception (ESP) would score higher on ESP tasks than nonbelievers. They found that not believing in ESP not only led to lower scores but also caused participants to behave contrarily. They also mentioned a previous study where a differential scoring effect was found. In this past study, those who believed in ESP scored progressively higher during the experiment, and those who did not believe in ESP showed a decline effect.

Murphy and Lester (1976) attempted to replicate a past study that found a correlation between the Myers-Briggs Type Indicator (MBTI) dimension of thinking-feeling and belief in ESP. They tested introductory psychology students and found a weak but significant positive correlation between belief in ESP and the feeling dimension. This may imply that psi is likely not something that we achieve with logical thought and attempting to do so causes it to elude us.

Tobacyk and Milford (1983) developed a 25-item questionnaire designed to assess belief in the paranormal. Using factor analysis, they found seven dimensions of paranormal belief: Traditional Religious Belief, Psi Belief, Witchcraft, Superstition, Spiritualism, Extraordinary (alien) Life Forms, and Precognition. While not all of the subscales on this questionnaire are of use in modern psi research, it seems that the dimensions of Traditional Religious Belief, Psi Belief, Spiritualism and Precognition could all be beneficial to someone that wants to conduct research on psi belief.

Haraldsson (1981) tested several variables for potential relationships with belief in psi, including religion, dream recall, politics, and survival after death (i.e. belief in an afterlife). He had students at the University of Iceland answer a "Sheep-Goat" questionnaire to determine the basis for their belief or disbelief in psi. There were significant positive correlations between belief in psi and life after death, religious beliefs, and dream interpretation, but most of the studies mentioned in the paper found the highest correlation between belief in psi and survival after death. Haraldsson asserted that there was a "moderate but apparently stable association between belief in psychical phenomena and belief in religion" (p. 297) based upon these findings.

Houtkooper and Haraldsson (1997) examined the correlation between guessing behavior and psi belief based on ten experiments carried out at the University of Iceland between 1977 and 1991. Participants completed a computerized ESP task in which they chose a target for each of 40 trials and were only given feedback on whether the choice was a hit or a miss. Houtkooper and Haraldsson found that there was clear evidence of test-retest reliability for participant patterns of guessing. These patterns included repetition avoidance (not choosing the same target multiple times), lateral preference ("asymmetry of calling," p. 122), central preference (tending to choose targets in the middle instead of ones on the ends), one-skipping repetition avoidance (identical to repetition avoidance, only each call is compared to the call made two before instead of one before), neighbor succession (choosing targets adjacent to the previous one in the row of targets), and short-term balancing (the number of sequences where all targets are called once each). The two most common cognitive heuristics used were repetition avoidance and central preference. Houtkooper and Haraldsson found that scoring remained relatively stable across participants for the ESP task, but they did not find support for the hypothesis that "sheep" (psi believers) tended to avoid repetition more than "goats" (psi disbelievers).

Watt and Nagtegaal (2000) examined the relationship between psi-mediated instrumental response (PMIR) and belief in good luck, expectations of success, and actual performance on a psi task. PMIR is a factor that pertains to nonintentional psi task performance. In order to measure PMIR, they used two tasks, one involving the United Kingdom (UK) National Lottery and the other using dice throws. PMIR was invoked by disincentivizing misses, or in other words punishing participants with a boring task if they did not score a hit on the psi task. Researchers believe that this paradigm motivates participants to tap into psi to find the correct target during a psi task. Watt and Nagtegaal found that PMIR and luck were significantly and positively correlated. There was a positive correlation between belief in good luck and expectations of success, but there was no difference in performance between participants based on their actual luckiness. On the other hand, Watt and Nagtegaal reported that participants who specifically believed their luck could affect their lottery success had significantly greater lottery success than other participants. This effect was not found for the dice-throwing task, however. They also did not find anything to suggest that their optimistic lottery players were being overconfident, and hypothesized that confidence of success varies depending on the type of task. They suggested that the large cash prize for winning the lottery may have been a powerful motivator that increased any psi effects that were present in the study.

Experimenter Effects

One explanation that has been proposed for psi effects in research is that they arise due to the beliefs of the experimenter. Time and again, the experimenter's level of belief is shown to have a considerable degree of influence on experimental results. While experimenter effects often manifest themselves in other areas of scientific inquiry, they seem to be especially salient in psi research due to the nonphysical nature of the effects that are being investigated. Many studies have been conducted on the role of experimenter effects in mental telepathy experiments, and a selection of them will now be presented and discussed.

One of the earliest studies on experimenter effects in psi research was carried out by Anderson and White (1937). They examined a series of ESP experiments done with American schoolchildren in which their teachers were responsible for administering the ESP tests, and found that the teacher's attitude toward the students had an effect on ESP scoring. They found that, overall, children who were more well-liked by their teachers tended to score better on the ESP tests than those who were not. Due to the age of this study, there may have been many methodological flaws in the data collection and analysis, but it is nonetheless notable for its historical significance.

Watt and Wiseman (2002) had two different experimenters carry out the same series of tests on 30 participants each. The tests used were a paranormal belief questionnaire, an ESP task and two cognitive ability tests, which were a syllogistic reasoning task (participants were asked to assess the validity of a conclusion based on two propositions that were asserted to be true) and Raven's Progressive Matrices (an intelligence test that measures abstract reasoning ability). A significant negative correlation was found for all participants between paranormal belief and syllogistic reasoning task performance. Using post-hoc analysis, they found that a significant negative experimenter effect existed on this measure for one of the two experimenters, but there were no experimenter effects on the ESP task.

Watt and Brady (2002) conducted two studies on the relationship between remote attention facilitation and experimenter effects. This type of psi task usually involves participants doing something boring (like staring at a lit candle) and pressing a button when their attention wavers. Typically, they will spend at least part of the experiment being "assisted" by someone in another room who attempts to help them stay focused on the candle, and then a computer records the number of button presses when they were being assisted and the number of presses when they were not. In this study, the experimenters included a variation of this task by dividing participants between an experimenter who had positive expectancies for the experiment and one who had negative expectancies. Due to an error in data collection, the first of the two studies could not be analyzed. In the second study, they found that remote attention facilitation had no effect on task performance, and there was no difference in psi scoring between positive and negative expectancy conditions. However, those who were tested by the positive expectancy experimenter gave significantly more positive evaluations of the experiment experience than those who were tested by the negative expectancy experimenter. These results may seem disappointing because no experimenter effect was found, but Watt revisited the experiment a year later and found more promising results, as described below.

Watt and Ramakers (2003) conducted a study similar to Watt & Brady's 2002 study of experimenter effects and remote attention facilitation, but added a twist: they used multiple experimenters who either believed in psi or did not (9 believers and 5 disbelievers each running 36 trials). They found that, overall, participants pressed the button fewer times when the experimenter was assisting them, and participants tested by believer experimenters had higher scores on the psi task, suggesting the presence of an experimenter effect. In other words, the total number of button presses when participants were being helped by a psi-believing experimenter was lower than in other conditions.

Wiseman and Schlitz (1997) examined the relationship between experimenter effects and participants' ability to detect when someone was staring at them. Participants answered a questionnaire prior to the experiment that assessed several factors including psi belief. The experimental runs were randomly split up into "stare" and "non-stare" trials where the experimenter stared at the participant or did not, and participants' electrodermal activity (EDA) was monitored throughout the experiment. The researchers found that the EDA was higher during stare trials for only one of the experimenters, which they believed may have been because that experimenter's participants reported a high belief in psi that just barely failed to reach significance.

Smith (2003) gave a questionnaire to 40 experimenters in an attempt to determine if they were "psi-conducive" (they tended to obtain significant results in their experiments) or "psi-inhibitory" (their experiments tended to fail to reach significance) and to identify psychological factors that might explain these trends. As part of the questionnaire, experimenters peer-rated each other on whether they were psi-conducive or psi-inhibitory, and the experimenters were categorized into one of the two groups based on the results. Measures included the Keirsey Temperament Sorter and a 6-item psi attitudes questionnaire. The experimenters also answered survey questions about their personal psi experiences and whether they had ever practiced a mental discipline (e.g. meditation). Smith found that psi-conduciveness was positively correlated with belief in one's own ESP ability, belief in one's own PK ability, belief that ESP is possible, and belief that ESP can be demonstrated in an experiment. These results suggest that psi researchers are most effective when they truly believe in psi, rather than approaching the topic with detached skepticism.

Luke, Delanoy and Sherwood (2008) examined the correlations among perceived luckiness, belief in luck, and belief in precognition. Participants completed a PMIR task in which the targets were images of fractals. A fractal is an abstract (typically computer generated) image that is sometimes used as the target in a psi experiment because it is ambiguous in form. 40 fractal images were created randomly with a fractal generator program and rated independently on how aesthetically pleasing they were by five judges, and no image was repeated in a given experimental run. The experimenters showed each participant four random fractals at a time from the pool of 40 images, and after making their selection one was randomly chosen in each trial to be the psi target, effectively making it a precognition task. Depending on their performance, the experimenters either rewarded the participants were with an erotic task (psi incentive) or punished them with a boring task (psi disincentive). The erotic task was much the same as the psi task, but erotic images were presented instead of fractals. Images were sorted by eight levels of eroticism and into three categories: male, female, and heterosexual couple. The eroticism level of the images was determined by how many hits they scored on the PMIR task (three hits corresponded to the lowest level, four hits was one level above that, etc.). They chose this design because of an earlier study by Stanford and Associates (1976) that stated that "contingent outcomes – in which degree of reward is related to degree of success – are preferable to dichotomous outcomes because they are theoretically and psychometrically more sensitive" (as cited in Luke, Delanoy and Sherwood, 2008). The unpleasant task also did not involve data collection, and required participants to observe a string of numbers and to press the left mouse key or the return key for every complete run of three consecutive odd numbers or even numbers. Participants worked on this task for four minutes, or less time if they scored one or two hits on the psi task. The experimenters found a significant precognition effect overall, as well as a significant positive correlation between luck scores on the questionnaire and precognitive performance. They also found significant correlations between psi task performance and belief in psi/belief in the paranormal, ($r(98) = .236$) and ($r(98) = .194$) respectively. Belief in psi was found to be a stronger predictor of psi task performance than belief in the paranormal. Erotic reactivity (a measure of how reactive a participant is to erotic stimuli) was also positively correlated with precognition performance.

Hitting/Missing Streaks

In psi research, a hitting streak refers to a string of correct guesses on a psi task, while a missing streak refers to a string of incorrect guesses. These are of interest when they deviate significantly from chance, because it means that research participants are harnessing psi to arrive at the correct answer, or in the case of a high psi-missing rate, they are actively avoiding it (for possible reasons that will be explored later). As hitting and missing streaks are paramount to this thesis' experimental analysis, research on this topic will now be presented and discussed.

Crandall and Hite (1983) discovered a psi-missing displacement effect (PMDE) in a pair of studies they conducted. That is, they found that participants who scored below mean chance expectation (MCE) on psi tasks scored significantly above chance when considering adjacent targets. They conducted four further studies to attempt to replicate this effect, two of which used Ganzfeld ESP (GESP) and clairvoyance tasks and two of which used a precognition task. In a typical ganzfeld experiment, a "receiver" is placed in a room relaxing in a comfortable chair with halved ping-pong balls over the eyes, having a red light shone on them. The receiver also wears a set of headphones through which white or pink noise (static) is played. The goal of this experimental method is to deprive the receiver of as many of their senses as possible, which researchers believe causes the mind to become overactive due to the lack of external stimuli (and therefore more likely to tap into psi). The receiver is typically kept in this state of sensory deprivation for half an hour, during which a "sender" observes a randomly chosen target and attempts to telepathically send this information to the receiver. The receiver speaks out loud during the 30 minutes, describing what he or she can see. Modern variations on Ganzfeld experiments may make use of special sensory deprivation tanks in which participants float on a saline solution, which has the added effect of depriving participants of their sense of touch. In their experiments, Crandall and Hite found strong PMDE effects for the task that each participant preferred, and in the precognition experiments, they found that PMDE occurred when participants were trying to guess the targets but not when they were trying to miss. They concluded that PMDE arose from improperly focused psi for those participants who psi-missed.

Crandall (1985) conducted further research on PMDE, in which he attempted to control task favorability so that he could determine how it influenced the degree of displacement. He assigned participants to two groups, one where the experimenter was cold and indifferent to the research (unfavorable condition), and the other where the experimenter was warm and enthusiastic about the research (favorable condition). He found that participants in the favorable condition showed a greater displacement effect and a larger number of psi-missers who scored above the mean chance expectation (MCE) for displacements. Participants in the unfavorable condition, on the other hand, had a greater number of psi-missers who were below MCE for displacement effects.

Kennedy (1979) examined consistent missing (CM) by reviewing the literature on the topic and analyzing experimental data for eleven participants. CM is defined as "the tendency of the subject in an ESP test to mistake particular symbols as certain other symbols..." (p. 113). Kennedy found that six of the eleven participants tested showed consistent CM effects in at least some of their data, and that CM effects were not consistently related to direct-hit scoring rate. CM effects were not found to be the main factor leading to negative scoring, and the factors that led to CM included similarities of meanings and associations between targets as well as visual similarities. Based on his results, Kennedy asserted that CM is one of the few proposed mechanisms for psi missing that does not require some kind of implied negative motivation, and called for further inquiry into this area of parapsychology due to its potentially vast significance for all psi-missing research.

Kelly (1980) examined consistent missing research more closely, and discovered that most prior investigations of CM chose target sets that were unrelated to CM in any meaningful way. He called for CM techniques to be deployed in a hypothesis-testing role, along with target sets specifically developed to assess for the presence of CM. He proposed two potential systematic effects that could be used to this end. The first of these was the reversal effect (when a participant envisions a psi target as a mirror image of itself), which Kelly tested by creating a deck of cards that each had one of the following letters: b, p, d, and q. This deck was used in a similar manner to a standard ESP card deck, and post-hoc analysis revealed that participants confused b and p most often, despite the fact that no actual CM effects were found. The other systematic effect Kelly proposed was the agent-percipient effect, which can be described briefly as different participants reliably interpreting the same stimulus in different ways. One example of this that Kelly mentioned is color-blindness tests, in which two people might see the same color splotch as two different colors. Finally, Kelly mentioned briefly that affective responses may play a vital role in the manifestation of psi effects, and that these may be a potential future topic of parapsychology research.

Palmer (2000) investigated psi-mediated instrumental response (PMIR) effects using a computer solitaire program. Participants played four games of solitaire each, at which point a computer entered their scores into a pseudorandom algorithm to determine whether they would be allowed to play additional games of solitaire. The experimenters also offered a $100 cash prize to the participant with the highest cumulative score, in order to incite psi effects. The dependent variable the researchers chose was the number of sets of solitaire completed, but no significant differences were found between participants on this measure. A post-hoc analysis conducted using the Solitaire Questionnaire that was given to all participants prior to testing found a significant correlation between participants' squared scores on the experiment and their scores on the questionnaire's addiction scale. Palmer asserted that this was evidence in support of his PMIR hypothesis, or in other words, that the most highly motivated participants would show the greatest psi effects.

Gambling and Psi

One setting in which psi effects may appear naturally is casinos. The games are played with real money at stake (which researchers believe incentivizes psi effects) and they commonly use random or pseudorandom processes to determine the outcomes. Many casino games rely on mathematical principles of probability, which makes them ideally suited to analyzing strings of hits and misses. Furthermore, video surveillance has allowed casinos to archive many hours of footage of various casino games, giving researchers an ample body of data on which to conduct probabilistic analyses. As my hypothesis hinges on strings of consecutive hits and misses and the probabilities of participants achieving various combinations of these, I believe that a thorough examination of the literature on probabilities in casino betting will inform my interpretation of the data. To be more precise, strings of consecutive hits and misses may be happening in the lab because of cognitive processes that parallel those used by gamblers to justify strings of hits and misses that appear to defy the rules of probability. Therefore, analyzing how gamblers use these cognitive processes could help to shed some light on the cognitive processes being used by participants in the EIU parapsychology lab.

Sundali and Croson (2006) examined patterns of betting in casino patrons during a span of 18 hours, using prerecorded footage of a roulette table. They were interested in finding out how often participants engaged in two common cognitive fallacies: the gambler's fallacy (a different outcome is more likely after a string of identical outcomes) and the hot hand fallacy (those who "hit" consistently are inherently more likely to continue to do so). They attempted to identify individuals who engaged in either of these fallacies based upon their betting patterns, and then to examine the intraparticipant correlation of these beliefs. They found that those who believed in the gambler's fallacy were more likely to also believe in the hot hand fallacy, and there was also a positive and significant interparticipant correlation between the two fallacies. They suggested locus of control as an underlying explanation for this result, but also mentioned that their study was limited because they had no way of contacting the gamblers to ask them why they engaged in these fallacies.

Fischer and Savranevski (2015) conducted a study in which participants were asked to predict the next several digits in a binary sequence and also explain the mechanism responsible for the pattern. The researchers' analysis of this task led them to propose four perceptual categories: two that were associated with the gambler's and hot hand fallacies (chance mechanisms and human performance), and two new ones called *periods and processes* and *traits and preferences*. Fischer and Savranevski then conducted a series of experiments investigating the relationships among these four categories and the alternation rates of the observed sequences under the categorical decisions structures of screening, discrimination, and classification. They found that chance mechanisms and periods and processes were stable across conditions, but the other two categories were context sensitive. Based on these results, they proposed that all of these categories represent cognitive structures that help us to identify and decode patterns in our physical and social environments.

Xu and Harvey (2014) examined hot hand effects in online gambling by analyzing 565,915 sports bets made by 776 online gamblers in 2010. Streaks were analyzed up to a maximum length of six. They found evidence for the gambler's fallacy in this population, because people who won were more likely to win again due to betting more conservatively, whereas those who lost were more likely to lose again due to betting more aggressively. By manipulating their odds of winning consistent with the gambler's fallacy, they created their own hot hand effects. It is worth mentioning that experiments carried out on online sports betting data are necessarily flawed due to the nature of online sports betting itself, as it is not representative of the random or pseudorandom conditions one observes in traditional casino games. Betters in this experiment were changing their behavior based on whether they were winning or losing, so it may be that participants in the EIU telepathy experiments changed their target selections based on whether the previous trial was a hit or a miss because they were engaging in similar behaviors. Such behavior could be a potential alternate explanation for scoring that is significantly above or below chance.

Nickerson and Butler (2009) examined participants' proficiency at producing random binary sequences. They did this by having participants come up with a string of outcomes on a hypothetical coin-tossing task, and found that they were generally ineffective at making the strings appear truly random. However, the experimenters cautioned that this should not be taken as evidence that participants are unable to generate random strings, because their efforts to create randomness should be judged with respect to their understanding of the nature of the task. It is not always clear how well participants understand the experimental task, so it is not possible to draw a definitive conclusion in this regard. They also mentioned that the experimental method precludes a demonstration of understanding of random processes in the first place. In other words, participants (and people in general) cannot create random strings, and even if they could, one could not be completely sure that the strings were actually random. It is possible, however, to create a pseudorandom string by relying on one of several different cognitive heuristics, and this is something that humans appear to do well. The overall results of this study suggest that the participants in the EIU parapsychology lab experiments cannot be engaging in truly random guessing because they are incapable of truly random guessing. However, they could still be guessing in a pseudorandom way that causes the results to appear random.

Evidence for Psi

There is a significant body of research that provides empirical support for psi phenomena. Since the 1950s, psychology researchers in many different disciplines have seen a considerable degree of improvement in experimental methodology as well as an increase in the rigor of analyses. These effects have been attributed to the creation of new technologies, such as computers that can ensure the same experimental protocol is done in exactly the same way, every time, as well as ensuring data are recorded and stored perfectly. Several psi researchers have called for a solution to the file drawer problem that plagues most areas of psychological research, suggesting that all psi research to be presented regardless of its statistical significance. As a result, some parapsychology journals (such as the Journal of Parapsychology) now encourage experimenters to submit reports on any experiment, no matter how significant the results are. Some of the published literature in support of psi's existence will now be reviewed.

Parker (2003) examined whether the question "Does psi exist?" is one that parapsychology researchers should be asking in the first place, and whether anyone should even be seeking an answer to it. He stated that it is agreed that some form of anomaly has been established, but nobody can agree on its nature. He called for the use of more adventurous experimental designs in developing new theories on psi. He cited digital autoganzfeld research (in which a computer randomly selects the psi targets to be displayed) as one example of a more adventurous experimental design, because it allows for the real-time recording of apparently high-quality psi as it enters consciousness. Finally, he warned that, regardless of whether psi turns out to be real or simply the result of fraudulent research, an answer to the question "Does psi exist?" may have unwanted implications for the field of parapsychology as a whole.

Bem and Honorton (1994) proposed that the replication rates and effect sizes achieved by using the ganzfeld procedure are statistically sufficient to serve as compelling evidence in support of psi. They supported their assertion by discussing some of the psi experiments that used a ganzfeld procedure and by reviewing meta-analyses conducted on older ganzfeld psi experiments. They found that, overall, replication rates were significantly above chance for the studies in the meta-analyses. For the series of experiments examined in their study, they chose to use dynamic psi targets (in this case, video clips) as opposed to the static targets that are typically used in psi experiments. They found that the use of video clips instead of static images led to significantly higher hit rates, and hypothesized that these findings were due to participants more readily using psi to focus on a dynamic target than a static one.

Standish et. al. (2003) conducted an experiment to determine whether participants could send and receive signals through an electromagnetically shielded room. The electromagnetic shielding was intended to help eliminate any kind of sensory leakage between the sender and receiver in the experiment. One participant was designated the receiver and one was designated the sender. The receiver was placed in an fMRI machine and instructed to attempt to receive information from the sender. The room in which the machine was set up had a layer of electromagnetic shielding around it, and the sender sat at a video monitor outside the room. On the sender's screen was a checkerboard pattern that either remained static or "flickered" (i.e., the squares would change colors), which they attempted to send to the receiver. Standish and colleagues found that one participant showed increased blood oxygenation in brain areas that are stimulated by the checkerboard pattern while the other did not. These findings suggest that psi effects work independently of electromagnetic fields, and can be used to activate certain areas of the receiver's brain based on what the sender is seeing. In other words, an information transfer may have taken place via mental telepathy, even if only on an unconscious level.

Nash (1976) conducted an experiment in which participants tried to correctly identify which of four paintings the experimenter was viewing during their experimental run. Participants underwent progressive relaxation for five minutes under the instruction of the experimenter, before the experimenter viewed the painting for 20 seconds and then showed the participants each of the four paintings in sequence. She found that participants psi-missed significantly above chance in this experiment, which she attributed to experimenter effects.

Kennedy (1995) examined the effectiveness of goal-oriented psi research. This research paradigm assumes that psi phenomena depend on a person's motivation for or benefit from the outcome of a random event, and do not depend on the complexity or information-processing aspects of the random process. He asserted that psi is goal-oriented, based on past experiments that have borne out this hypothesis which he discussed in this article. However, he also mentioned that goal-oriented psi effects are not always statistically supported, especially when the goals are too broad instead of being small, specific, and testable. As such, he chose to focus on redundant opportunities for psi to operate (e.g. random number generator (RNG) or voting procedures) because this is the form of goal-oriented psi that is most significantly suited for statistical analysis and also has strong empirical support. Based on his review of the experiments cited in this article, he concluded that majority-vote studies are the most efficient way to examine goal-oriented psi because these data provide a basis for investigating goal-oriented psi in many other contexts and at all hierarchical levels.

Blackmore and Rose (1997) hypothesized that people are more likely to misconstrue confusions between reality and imagination as psi effects, either because they believe in psi or because such mental states are conducive to psi effects. They tested these hypotheses by conducting an experiment in which participants were shown slides of various objects and asked to imagine other objects, with the goal of creating the type of confusion they believed may be conducive to psi. Participants also completed the Belief in the Paranormal scale to assess whether or not they believed in psi. After three weeks, they were questioned on whether they had seen the objects or imagined them, with half of the imagined objects being randomly chosen as psi targets. They did not find any correlation between the number of false memories and the Belief in the Paranormal score. There were significantly more false memories on target objects than nontarget objects. The experimenters considered these findings to be evidence in support of the reality-imagination confusion being psi-conducive. In other words, people may be able to tap into psi more easily if their reality testing is flawed than if they remain in touch with reality at all times. This may also serve as an explanation for why people seem to tap into psi more often while experiencing altered states of consciousness (e.g. drug trips and religious experiences).

Sheldrake (2014) examined a phenomenon he termed “telephone telepathy”: the relatively common occurrence of thinking about someone right before they call you or send you a text, or knowing that they are trying to reach you when the phone rings before you even answer it. He claimed that it usually occurs between people who have strong bonds or emotional connections with each other. He tested this hypothesis by conducting an experiment in which participants were asked to identify one of four people who was calling them before picking up the phone, and found that the average scores were significantly above the 25% hit rate expected by chance. He tested participants in the lab under video surveillance at first, but later switched to an automated system that allowed experimental data to be collected through the use of a cell phone application. In both experimental conditions Sheldrake found strong support for the existence of telephone telepathy. He also mentioned experiments that tested this phenomenon with emails and SMS messages, all of which reported hit rates that were significantly above chance. Psi effects appear to be more common between people that share a strong emotional bond, and the results of these experiments provide strong support both for the existence of psi and for close relationships between people being a contributing factor towards the occurrence of psi effects.

Skeptics’ Viewpoints on Psi Phenomena

Skeptics have forwarded valid counterpoints to the claims put forth by proponents of parapsychology, such as that the effect sizes of parapsychology experiments have declined with increasing experimental rigor, or that many psi experimenters are unable to replicate findings even when using an identical experimental design.

West (2012) examined some of the questions that have been raised about flaws in parapsychology research. He mentioned that effect sizes in parapsychology have declined as experimental methods have become more rigorous, but pointed out that skeptics' claims of sensory leakage in the experiments have been rebuffed. This is one of several alternative explanations for the presence of psi effects that skeptics like to propose, but rarely if ever does the experimental methodology allow for such a thing to occur, as psi experimenters often take great pains to ensure that their participants cannot see or hear each other. West mentioned that psi experimenters were unable to agree on the criteria for psi to manifest in the lab. He also mentioned that reports of real-world psi have become increasingly rare in parapsychology journals, despite the fact that they are as prevalent now as they were 100 years ago. He attributed this to a hesitation on the part of psi researchers to publish findings that are often qualitative or even anecdotal in nature, because these sources are viewed as less scientific than a controlled study conducted in a lab.

Greenhouse (1991) considered the basis for skeptics' resistance to accepting the existence of parapsychology as a legitimate field of research. He cited the lack of a parapsychological model as one of the biggest factors in skeptics' continued resistance, as psi effects contradict much of what we know about physics and biology, and thus most scientists are reluctant to accept it on this basis. He mentioned that experimenter expectancy is an important factor to consider in all psi experiments, and underestimating it could be a major error. This claim echoes the findings of many parapsychology researchers that have examined experimenter effects and found them to be significant (see the relevant literature review section above for more detail). Greenhouse concluded his essay by discussing the role of meta-analysis in parapsychology research. Skeptics have claimed that the use of meta-analysis to find significant results across experiments that were not individually significant is little more than wishful thinking; that parapsychology researchers are twisting the data around to suit their hypotheses. Greenhouse mentioned that this is indeed the goal of meta-analysis; however, researchers can also use it as a tool to examine the effects of differences in experimental design as well as the effects of moderating variables. He suggested that these applications are more valid uses of meta-analysis that parapsychology researchers rarely utilize, and that the field could benefit from applying it more broadly.

Findings from Eastern Illinois University Parapsychology Lab

The current thesis involves reanalysis of archival data from the EIU parapsychology lab with the goal of identifying patterns or strings of hits and misses. The goal of this section is to summarize the methodology and findings of previous studies conducted in the lab, including theses completed by Warner-Angel (2016); Bunfill (2006); McWhorter (2005); Yarnall (2003); Conner (2009); and Parker (2006).

Warner-Angel (2016) outlined an example of a typical methodology used in the parapsychology lab experiments. She stated that

> A receiver, picked at random from a group of senders, sits in an isolated room in front of a foam board marked with four target circles (one in each corner). Three red discs with strips of Velcro adhered to the backs are used as target markers. The receiver is viewed via closed video feed by a group of senders who are facing an identical foam board in a separate room with only two red discs, one marked (+$) and the other marked (-$). Target locations for the (+$) and (-$) discs are selected in the sender room by using random numbers. The receiver is signaled by a bell tone that the target locations have been chosen. The senders try to transmit the (+$) target location to the receiver telepathically and the receiver attempts to locate it marking it on the foam board with one of the discs. The receiver continues to place discs on the board until either the (+$) or (-$) target is found. If the (+$) is selected a bell is sounded to indicate that the receiver has scored a "hit." If the (-$) target is chosen, a buzzer is sounded to indicate that the receiver has scored a "miss." Several trials (usually 16) are run. The receiver is returned to the sender room and another receiver is randomly selected from the group. A single response is referred to as a trial, a set of trials constitutes a run, a series of runs constitutes a session, and several sessions comprise an experiment. (Warner-Angel, 2016, pp. 18-19)

Warner-Angel's (2016) thesis analyzed data collected in the spring semester of 2011. She hypothesized that monetary incentive and the presence of competition among participants would significantly increase hit rates across all participants. Her results were not significant initially, but when separating participants with high belief scores from participants with low belief scores she found a strong effect, with high belief participants (a 5, 6, or 7 on a 7-point Likert scale) scoring significantly higher than low belief participants (a 1, 2 or 3 on a 7-point Likert scale), ($t(70) = -3.16$, $p < .001$). These findings replicated the findings of McWhorter (2005) who used a similar methodology, and who also found a direct positive correlation between belief in psi and hit rates.

Yarnall (2003) reported on an experiment using the basic methodology described by Warner-Angel, with the added variable of dividing participants into an "in-group" of participants who met and chatted with the senders and experimenters prior to the experiment, and an "out-group" of participants who were isolated in another room prior to the experiment. Yarnall hypothesized that the participants in the in-group would score significantly above chance due to a psi-conduciveness effect, while those in the out-group would score significantly below chance due to a natural tendency toward a psi-inhibitive mindset. Yarnall found that the in-group scored higher than the out-group during all twelve data collection sessions ($t = 3.84$, $p < .001$, two-tailed).

Bunfill (2006) examined the effects of belief in psi, decline effects, and practice effects. He hypothesized that participants who had been involved in a prior experiment in the EIU lab would start off scoring higher than those who had not been involved in a prior experiment, but their scores would decline over the course of the session due to boredom or fatigue. He found that there was indeed a significant decline effect in the data, and while psi belief and scoring were not significantly correlated, psi belief had a mediating effect on decline rates, with high-belief participants showing a slower decline than low-belief participants.

Conner (2009) examined a series of three experiments conducted in the spring of 2006. She was primarily interested in determining which types of participant groups are psi-conducive and which types are psi-inhibitory, based on a range of variables. She found a significant positive correlation between groups of participants rated as friendlier by research assistants and hit rates ($r(7) = .73$, $p = .027$), replicating the findings of Yarnall (2003). She hypothesized that specific subgroups would miss targets above probability when instructions to hit were given and would hit targets above probability when instructed to miss.

Conner found that men with low belief in psi tended to achieve significantly higher scores when instructed to miss the target than men with high belief in psi, but this difference was not found for women. She also found that, in general, participants with a high belief in telepathy tended to score lower when instructed to miss than those with low or medium belief. Conner theorized that high belief men who missed above chance when instructed to miss were following the experimenter's directions, while low belief men were acting contrarily due to some subconscious psychological factor. The last stage of Conner's analysis focused on experiments in which senders and receivers were grouped by sex, so that men sent to women and vice versa in two experimental conditions, men sent to men and women sent to women in the other two conditions. She found that, overall, women achieved significantly higher hit rates than men regardless of the condition.

In her discussion section, Conner suggested that analyses based on friendliness ratings are important because if the above chance hit rates of friendly groups (psi-hitting) are combined with the below chance hit rates of less friendly groups (psi-missing), the results cancel each other out and appear to be random. Conner suggested that there may be an experimenter effect present in her data, because participants who are enrolled in one of Gruber's classes generally score differently in his psi experiments than participants who are not enrolled in a class with him. This is especially true of men, for reasons that Conner admitted she does not fully understand.

Regarding Conner's discussion of findings from the EIU parapsychology studies, she mentioned that Gruber and colleagues have speculated that "social fields" may have an effect on psi scoring, and they have designed several experiments to test this hypothesis. Social fields are a type of psi-mediating variable that is believed to strengthen or create psi effects in the lab; in this case, Conner was specifically interested in finding this effect between participants who were more friendly and thus were believed to more easily create social fields in the lab.

Parker (2006) analyzed the effects of order within sessions (whether a participant in a given session went first, last, or in between) and belief in telepathy on telepathy task performance. In the second stage of his analysis, he compared participants who had also participated in the first stage of the experiment with those who did not to determine whether there was a practice effect. The experiments he analyzed were similar to those that Conner analyzed, although no attempt was made to control for friendliness or participants' contact with the experimenter outside of the lab setting. He hypothesized that participants who went first and last in a given experimental session would have lower scores than other participants, those with high belief in telepathy would score better than those with low belief in telepathy, and those who participated in the experiment more than once would show practice effects. He found that participants who went first and second scored significantly below chance (and the last person run), and there was also a short break between the fifth and sixth runs that may have disrupted the participants' focus. He found that participants who reported a high belief in telepathy scored below the rate expected by chance, but no effect was found for participants with low belief in telepathy. Overall, the data supported the existence of the practice effect that Parker hypothesized, with participants who had been in both stages of the experiment scoring significantly higher in the second stage than those who had not ($t(296) = 2.11$, $p = .035$).

Regarding his findings, Parker speculated that the order effect for participants who went first and second (known as a "first person effect" in the literature) may have been due to being unfamiliar with the methodology, which may have also caused them to be defensive or nervous. The last participant's low scores were thought to be due to fatigue or boredom. He speculated that the sex effect may have been due to women being more sensitive to psi information and therefore more likely to defensively reject it initially.

Current Study

The current study was designed to analyze strings of hits and strings of misses in a series of telepathy experiments conducted in the EIU parapsychology lab. The basic question I intended to answer was: Are participants more likely to score a hit after a hit and a miss after a miss? In other words, are there more strings of hits and strings of misses than we would expect by chance? The analysis involved examining the percentage of hits followed by hits, and misses followed by misses ("strings"), along with extended strings of hits and extended strings of misses. The data were also split by belief and explored for strings. My hypotheses are as follows: There will be more strings of hits and/or strings of misses than expected by chance, there will be a difference between the number of strings of hits and the number of strings of misses, and the participants' scores on a 7-point belief in psi scale will be related to the frequency of strings of hits and/or strings of misses. If there are more strings than what would be expected by chance, then it could indicate the presence of psi phenomena. Gruber proposed the following explanation for psi hitting streaks: "Participants are in a state of mind where they are getting better information, or after getting a hit, it creates a state of mind where they are more likely to get another hit." (Gruber, personal communication, March 10, 2017) If participants show above chance strings of misses, then this might provide support for the existence of some kind of underlying phenomenon that causes participants to actively avoid choosing the correct target. Psi-missing researchers such as Kennedy (1979, 1994, 1995) and Parker (2006) would likely take such a finding to be compelling support for their theories. Finding above chance strings of hits or misses has interesting implications and may indicate that the gambler's fallacy is false. In other words, gamblers may accurately perceive above chance strings of winning and losing.

The data to be reanalyzed in this study were chosen from methodologies that used both hit and miss targets (versus only a single hit target used for many years in the lab) allowing for the computation of strings of hits and strings of misses. The use of both hit and miss targets was introduced in the fall of 2010. Data were also chosen based on methodologies that allowed for direct feedback to receivers following hits or misses (versus the simultaneous testing of multiple participants without providing feedback) and use of monetary incentive. The Warner-Angel (set A) and Hinman (set B) data were chosen based on these criteria and have already been shown to produce highly significant findings. Two other data sets (C and D) were chosen using the same criteria, but did not produce significant findings on previous analysis. These datasets were reanalyzed to determine if they might show strings of hits or misses at above chance rates.

Method

The methodology involved reanalyzing past datasets from mental telepathy experiments conducted at EIU. The probabilistic chance of a hit following a hit is 25%, the chance of a miss following a miss is 25%, and there is a 50% chance of a hit-miss or miss-hit pairing. The experiments that were analyzed included strict controls that were designed to ensure that participants were not receiving or sending information through any of their five senses. Senders and receivers in the experiments were isolated in two different rooms located on different floors of the Physical Sciences building, and the experimenters tested the rooms to make sure that no sensory "leaks" were occurring (for instance, participants being able to hear each other through the building's air ducts). The sender could view the receiver through a video feed -- the receiver was not able to see or hear the sender. The receiver was seated in front of a foam board (or a website designed for the task) marked with four target circles (one in each corner), and was instructed to place a marker on the board to indicate which of the four circles contained the correct target. The researchers signaled the receiver with a bell for a hit and a buzzer for a miss. Each receiver completed 16 experimental trials in this fashion, which constitutes one experimental run.

The data were selected from four semesters: Fall 2010, Spring 2011, Fall 2011, and Fall 2017. The semesters were chosen as they employed both hit and miss targets with feedback provided. The two nonsignificant datasets were also chosen to determine if strings of hits and misses appeared significantly above the chance expectation, as this type of analysis was never conducted on these datasets originally. It is interesting to note that the participants in the studies that did not produce significant findings were not enrolled in the lead researcher's introductory psychology class. Below is a brief description of the methods from the datasets with significant findings, as written up by Hinman and Warner-Angel in their master's theses. The quotations below are condensed from the abstracts of their theses. Hinman (2017) wrote

> The current study was designed to explore how the variables of belief, gender, and experimenter effects regarding setting influence participants' results on a psi task. Sender and receiver groups were divided by gender. It was hypothesized that gender, belief, and the social atmosphere of the experiment had the potential to increase either psi-hitting or psi-missing. Results indicate that...far more participants than expected by chance produced extreme scores (scores of 4 or 12 hits out of 16 trials at $p < .002$). Only female receivers were run on five of the six data collection sessions, and all extreme scores observed happened on these sessions. The odds that this occurred strictly due to chance are calculated at $p = .0003$.

Warner-Angel (2017) wrote

> The current study was designed to explore variables influencing psi performance including gender, belief in psi, and competition between men and women. [A key] feature of the experiment involved... monetary incentive to increase competition in a game-show type of environment.... [with the goal of] exploring

> the role of competition between groups of senders and receivers as a facilitator of telepathic communication.... It was hypothesized that [mental telepathy] belief ratings would correspond with hit rates—specifically that participants with high belief in psi phenomena would score above chance while those with low belief scores would score below chance. The results of an independent samples t-test show that high belief participants (M = 8.92, SD= 1.75) scored significantly higher on the telepathy task than low belief participants (M = 7.67, SD= 1.53), $t(70)= -3.16$, $p < .001$ (one-tailed). Given the strength of these findings, it appears that psi phenomena and mental telepathy, specifically, do occur and may be influenced by belief in psi, group setting, and competition.

In order to carry out the analysis, pairs (strings of two hits or misses) and triplets (strings of three hits or misses) were obtained for each run in the dataset. Each participant completed a run of 16 trials, with a mean expected hit rate of 8 trials (50%). Pairs were coded as follows: pair 1 consists of the first trial and the second trial, pair 2 consists of the second and third trials, etc. Triplets were coded similarly, with triplet 1 consisting of the first 3 trials, triplet 2 consisting of trials 2 through 4, etc. Because pairs and triplets were coded in this way, a run of 16 trials could contain up to 15 pairs and 14 triplets. The total number of pairs of hits and misses was computed in SPSS. Strings of four or more hits and misses (up to ten) were counted manually. Statistical frequencies were then obtained for trials 1 through 16, pairs 1 through 15, triplets 1 through 14, and extended strings. The file was then split by participant belief, which describes how each participant scored on a 7-point belief in psi scale administered prior to the experiment, with 1-3 indicating low belief, 5-7 indicating high belief, and 4 indicating "middle" belief. A series of t tests were used to compare individual trial data (1-16). All other results were analyzed using chi-squares or binomial calculations because these data were not interval or ratio data. Finally, an analysis investigating the presence of an order effect in the data was carried out by grouping runs into thirds: a beginning, middle, and end. The data were split into thirds as follows: Trial 1 through 5, 6 through 11, and 12 through 16; pair 1 through 5, 6 through 10, and 11 through 15; and triplet 1 through 5, 6 through 9, and 10 through 14.

Results

It was hypothesized that there would be more strings of hits and/or misses than expected by chance (A); there would be a difference between the number of strings of hits and the number of strings of misses (B); and participant scores on a 7-point belief in psi scale would be related to the frequency of hits and misses (C).

Across all four datasets, there were 348 participants that ran a total of 5568 trials. We would expect to find 2784 hits and 2784 misses (50% each) by chance. Participants scored a total of 2812 hits (M = 8.08 per run) and a total of 2755 misses (M = 7.92), which according to binomial probability did not differ from chance expectation. For totals of hits and misses obtained in each dataset and across all datasets, see Table 1.

With regard to pairs of hits and misses, each participant could obtain 15 pairs per run of 16 trials for a total of 5220 possible pairs. There was a mean chance expectation (MCE) of 1305 pairs of hits and 1305 pairs of misses (25% each, $M = 3.75$), with the additional 50% comprised of one hit and one miss. Participants scored a total of 1310 pairs of hits ($M = 3.76$) and a total of 1270 pairs of misses ($M = 3.65$), which according to binomial probability did not differ from chance expectation. For totals of pairs of hits and misses obtained in each dataset and across datasets, see Table 1.

Regarding strings of three (triplets), each participant could obtain 14 triplets per run of 16 trials for a total of 4872 possible triplets. There was a mean chance expectation of 609 (12.5% each, $M = 1.75$) for strings of hits and strings of misses. Participants scored a total of 602 triplets of hits ($M = 1.73$) and a total of 584 triplets of misses ($M = 1.68$), which is not significant. For totals of triplets of hits and misses obtained in each dataset and across datasets, see Table 1. As none of the above findings were significant, hypotheses A and B were not confirmed.

The combined datasets were then split by belief to determine if participants with high belief scores and low belief scores had different amounts of strings of hits and misses. The participants with scores of 1, 2, and 3 on a 7 point scale were determined to be low belief (N = 154); those with scores of 5, 6, and 7 were determined to be high belief (N = 141); while those with scores of 4 were determined to be middle belief (N = 51). The total number of hits for high belief participants was 1176 ($M = 8.33$ per run), while the total number of misses was 1080 ($M = 7.66$). The total number of hits for low belief participants was 1212 ($M = 7.86$), while the total number of misses was 1251 ($M = 8.12$). A t test for independent means showed a significant difference between hits for high belief compared with hits for low belief participants resulting in $t = -2.08$, $p = .039$. For combined totals of hits and misses across all datasets, see Table 2 and Figure 1.

The total number of pairs of hits for high belief participants was 556 ($M = 3.94$ per run), while the total number of pairs of misses was 464 ($M = 3.29$). The total number of pairs of hits for low belief participants was 549 ($M = 3.56$), while the total number of pairs of misses was 601 ($M = 3.90$). As strings of hits and misses are not interval or ratio data, chi-square tests for independence were used to analyze the pairs, triplets and extended strings. A chi-square test indicated a significant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 9.915$ (1, N = 295), $p = .0016$. For combined totals of pairs of hits and misses across all datasets, see Table 3 and Figure 2.

The total number of triplets of hits for high belief participants was 265 ($M = 1.88$ per run), while the total number of triplets of misses was 199 ($M = 1.41$). The total number of triplets of hits for low belief participants was 236 ($M = 1.53$), while the total number of triplets of misses was 289 ($M = 1.88$). A chi-square test indicated a significant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 14.57$ (1, N = 295), $p = .00014$. For combined totals of triplets of hits and misses across all datasets, see Table 4 and Figure 3.

Table 1

Descriptive Statistics for Individual and Combined Datasets

	N	Trials	Total hit	Total miss	pair miss	pair hit	Total pair	triple miss	triple hit	Total triple
Dataset A	98	1568	804	764	337	380	717	151	183	334
Dataset B	53	848	418	430	205	193	398	103	94	197
Dataset C	90	1440	732	707	336	349	685	159	161	320
Dataset D	107	1712	858	854	392	388	780	171	164	335
Total	348	5568	2812	2755	1270	1310	2580	584	602	1186

Combined datasets split by belief

Table 2

Combined trial scores across all datasets split by belief

	Trial Scores (LB N = 154; HB N = 141)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Trial 1	150	145	80 (51.9)	74 (48.1)	70 (49.6)	71 (50.4)
Trial 2	136	159	73 (47.4)	81 (52.6)	63 (44.7)	81 (57.4)
Trial 3	143	152	76 (49.4)	78 (50.6)	67 (47.5)	74 (52.5)
Trial 4	153	142	76 (49.4)	78 (50.6)	77 (54.6)	64 (45.4)

Trial 5	147	148	69 (44.8)	85 (55.2)	78 (55.3)	63 (44.7)
Trial 6	161	134	79 (51.3)	75 (48.7)	82 (58.2)	59 (41.8)
Trial 7	149	146	77 (50.0)	77 (50.0)	72 (51.1)	69 (48.9)
Trial 8	146	149	62 (40.3)	92 (59.7)	84 (59.6)	57 (40.4)
Trial 9	144	151	71 (46.1)	83 (53.9)	73 (51.8)	68 (48.2)
Trial 10	142	153	84 (54.5)	70 (45.5)	58 (41.1)	83 (58.9)
Trial 11	149	146	81 (52.6)	73 (47.4)	68 (48.2)	73 (51.8)
Trial 12	146	149	67 (43.5)	87 (56.5)	79 (56.0)	62 (44.0)
Trial 13	146	149	73 (47.4)	81 (52.6)	73 (51.8)	68 (48.2)
Trial 14	156	139	79 (51.3)	75 (48.7)	77 (54.6)	64 (45.4)
Trial 15	159	136	78 (50.6)	76 (49.4)	81 (57.5)	60 (42.5)
Trial 16	161	133	87 (56.5)	66 (42.9)	74 (52.5)	67 (47.5)
Total	2388	2331	1212 (49.2)	1251 (50.8)	1176 (52.1)	1080 (47.9)

Note: LB = low belief; HB = high belief

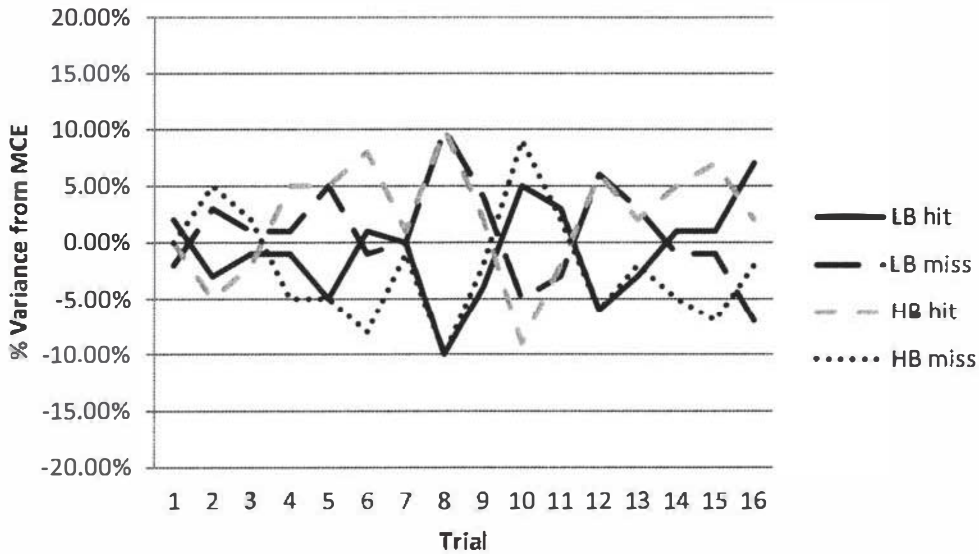

Figure 1. Combined scores for each trial split by belief.

Table 3

Combined pair scores across all datasets split by belief

			Pair Scores (LB N = 154; HB N = 141)			
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Pair 1	72	81	42 (27.3)	43 (27.9)	30 (21.3)	38 (27.0)
Pair 2	59	75	34 (22.1)	39 (25.3)	25 (17.7)	36 (25.5)
Pair 3	75	74	38 (24.7)	40 (26.0)	37 (26.2)	34 (24.1)
Pair 4	72	67	36 (23.4)	45 (29.2)	36 (25.5)	22 (15.6)
Pair 5	76	63	34 (22.1)	40 (26.0)	42 (29.8)	23 (16.3)
Pair 6	87	72	41 (26.6)	39 (25.3)	46 (32.6)	33 (23.4)

Pair 7	70	70	30 (19.5)	45 (29.2)	40 (28.4)	25 (17.7)
Pair 8	75	80	32 (20.8)	53 (34.4)	43 (30.5)	27 (19.1)
Pair 9	67	76	40 (26.0)	39 (25.3)	27 (19.1)	37 (26.2)
Pair 10	72	76	49 (31.8)	38 (24.7)	23 (16.3)	38 (27.0)
Pair 11	68	68	31 (20.1)	37 (24.0)	37 (26.2)	31 (22.0)
Pair 12	69	73	29 (18.8)	43 (27.9)	41 (29.1)	30 (21.3)
Pair 13	76	69	36 (23.4)	38 (24.7)	40 (28.4)	31 (22.0)
Pair 14	84	64	36 (23.4)	33 (21.4)	48 (34.0)	31 (22.0)
Pair 15	83	57	41 (26.6)	29 (18.8)	42 (29.8)	28 (19.9)
Total	1105	1065	549 (23.8)	601 (26.0)	556 (26.3)	464 (21.9)

Note: LB = low belief; HB = high belief

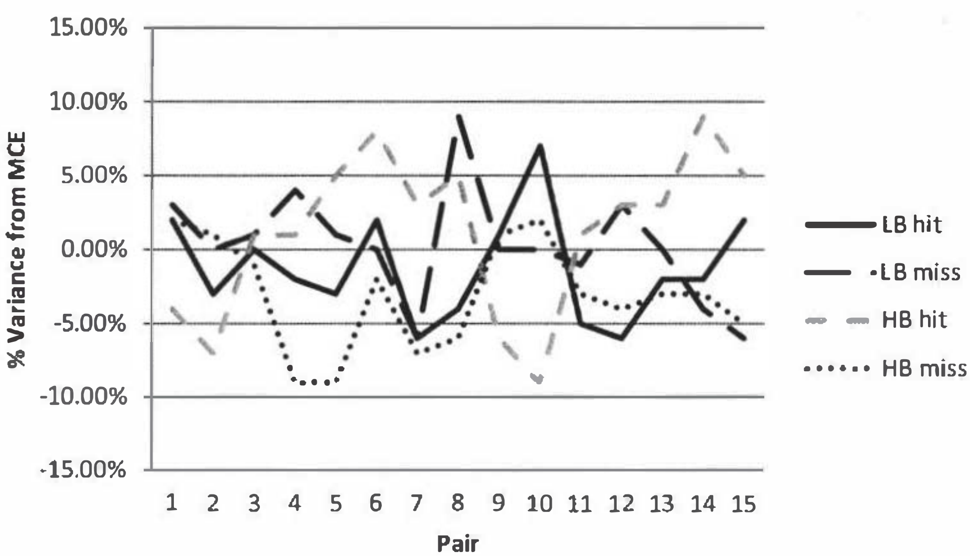

Figure 2. Combined scores of pairs for each trial split by belief.

Table 4

Combined triplet scores across all datasets split by belief

	Triple Scores (LB N = 154; HB N = 141)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Triple 1	35	39	22 (14.3)	22 (14.3)	13 (9.2)	17 (12.1)
Triple 2	29	36	16 (10.4)	20 (13.0)	13 (9.2)	16 (11.3)
Triple 3	39	39	19 (12.3)	27 (17.5)	20 (14.2)	12 (8.5)
Triple 4	38	30	19 (12.3)	21 (13.6)	19 (13.5)	9 (6.4)
Triple 5	40	33	20 (13.0)	18 (11.7)	20 (14.2)	15 (10.6)
Triple 6	41	34	15 (9.7)	22 (14.3)	26 (18.4)	12 (8.5)
Triple 7	35	41	14 (9.1)	30 (19.5)	21 (14.9)	11 (7.8)
Triple 8	32	37	17 (11.0)	22 (14.3)	15 (10.6)	15 (10.6)
Triple 9	34	41	23 (14.9)	21 (13.6)	11 (7.8)	20 (14.2)
Triple 10	29	35	15 (9.7)	19 (12.3)	14 (9.9)	16 (11.3)
Triple 11	35	34	12 (7.8)	18 (11.7)	23 (16.3)	16 (11.3)
Triple 12	36	33	13 (8.4)	20 (13.0)	23 (16.3)	13 (9.2)
Triple 13	36	32	14 (9.1)	17 (11.0)	22 (15.6)	15 (10.6)
Triple 14	42	24	17 (11.0)	12 (7.8)	25 (17.7)	12 (8.5)
Total	501	488	236 (10.9)	289 (13.4)	265 (13.4)	199 (10.1)

Note: LB = low belief; HB = high belief

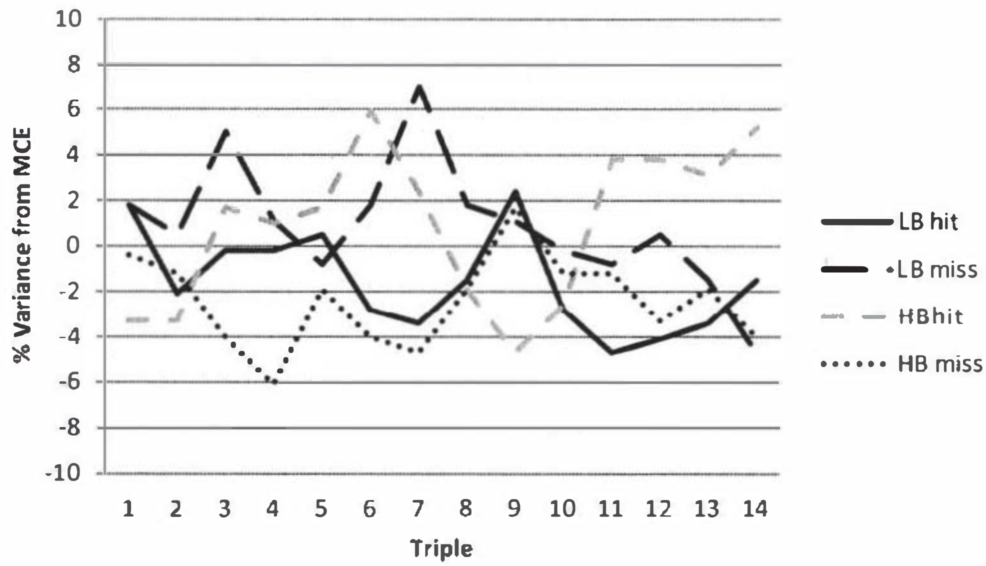

Figure 3. Combined triplet scores across all datasets split by belief.

As part of the data analysis, the datasets were examined for strings of up to ten hits and misses. Frequencies were then obtained for these extended strings. These results are presented in Table 5 below while the extended strings for each dataset can be found in appendix A. It is worth mentioning that there is overlap in the counting of the extended strings, as it would be impossible to obtain them otherwise (a string of five must contain a string of four, for instance). While relatively rare, multiple strings of four and five did occur in some runs, these have simply been counted as two strings of four or five.

The total number of quad (4) hits for high belief participants was 60 (*M* = 0.43), while the total number of quad misses was 43 (*M* = 0.31). The total number of quad hits for low belief participants was 54 (*M* = 0.35), while the total number of quad misses was 65 (*M* = 0.42). A chi-square test indicated a close to significant relationship between hits and misses when comparing high and low belief participants, $\chi^2 = 3.66$ (1, N = 295), $p =$ .055. For combined totals of quads across all datasets, see Table 5.

The total number of quintuple (5) hits for high belief participants was 30 ($M =$ 0.21), while the total number of quintuple misses was 14 ($M = 0.09$). The total number of quintuple hits for low belief participants was 28 ($M = 0.18$), while the total number of quintuple misses was 34 ($M = 0.22$). A chi-square test indicated a significant relationship between hits and misses when comparing high and low belief participants, $\chi^2 = 5.504$ (1, N = 295), $p = .019$. For combined totals of quintuples across all datasets, see Table 5. Further extended strings were not analyzed due to small sample sizes.

Table 5

Extended strings combined across all datasets

		Extended strings (LB N = 154; HB N = 141; MB N = 51)		
	Frequency	LB (%)	HB (%)	MB (%)
4hit	133	54 (35.1)	60 (42.6)	19 (37.3)
4miss	128	65 (42.2)	43 (30.5)	20 (39.2)
5hit	68	28 (18.2)	30 (21.3)	10 (19.6)
5miss	57	34 (22.1)	14 (9.9)	9 (17.6)
6hit	29	10 (6.5)	14 (9.9)	5 (9.8)
6miss	25	13 (8.4)	6 (4.3)	6 (11.8)
7hit	9	3 (1.9)	4 (2.8)	2 (3.9)
7miss	14	7 (4.5)	4 (2.8)	3 (5.9)
8hit	6	2 (1.3)	2 (1.4)	2 (3.9)
8miss	7	4 (2.6)	2 (1.4)	1 (1.9)
9hit	4	1 (0.6)	1 (0.7)	2 (3.9)
9miss	6	3 (1.9)	2 (1.4)	1 (1.9)
10hit	2	1 (0.6)	0 (0)	1 (1.9)
10miss	1	0 (0)	1 (0.7)	0 (0)

Note: LB = low belief; HB = high belief; MB = middle belief

Position Effects

A visual inspection of the graphs suggested a position effect (hit and miss rates varying consistently over the course of experimental runs). In order to examine this phenomenon, the trials were divided into segments. Trials were divided into three groupings, with 1 through 5 being the first third (5 trials), 6 through 11 being the second third (6 trials), and 12 through 16 being the third third (5 trials). Pairs were split similarly, with 1 through 5 being the first third (5 pairs), 6 through 10 being the second third (5 pairs), and 11 through 15 being the third third (5 pairs). Triplets were split into triplets 1 through 5 being the first third (5 triplets), triplets 6 through 9 being the second third (4 triplets), and triplets 10 through 14 being the third third (5 triplets). The split datasets were graphed and visually inspected for the presence of a position effect, based on this it was determined that the third segment showed the greatest variance. Post-hoc chi-squares were run for the third segment of the trials, pairs, and triplets.

The total number of hits for high belief participants on trials 12 through 16 was 384 (M = 2.72), while the total number of misses on trials 12 through 16 was 321 (M = 2.28). The total number of hits for low belief participants on trials 12 through 16 was 384 (M = 2.49), while the total number of misses on trials 12 through 16 was 385 (M = 2.5). A chi-square test indicated no significant relationship between hit and miss rates when comparing high and low belief participants, χ^2 = 3.0286 (1, N = 295), p = .082. For totals of all datasets split by belief and grouped into thirds, see Table 6 and Figure 4. While individual trials did not show a significant end of run order effect, there were significant differences between high and low belief participants for strings of hits and misses.

The total number of hits for high belief participants on pairs 11 through 15 was 207 (*M* = 1.47), while the total number of misses on pairs 11 through 15 was 151 (*M* = 1.07). The total number of hits for low belief participants on pairs 11 through 15 was 173 (*M* = 1.12), while the total number of misses on pairs 11 through 15 was 180 (*M* = 1.17). A chi-square test indicated a significant relationship between belief and hit rates when comparing high and low belief participants, χ^2 = 5.548 (1, N = 295), *p* = .018. For totals of all datasets split by belief and grouped into thirds, see Table 6 and Figure 5.

The total number of hits for high belief participants on triplets 10 through 14 was 107 (*M* = 0.76), while the total number of misses on triplets 10 through 14 was 72 (*M* = 0.51). The total number of hits for low belief participants on triplets 10 through 14 was 71 (*M* = 0.46), while the total number of misses on triplets 10 through 14 was 86 (*M* = 0.56). A chi-square test indicated a significant relationship between belief and hit rates when comparing high and low belief participants, χ^2 = 7.1114 (1, N = 295), *p* = .0076. For totals of all datasets split by belief and grouped into thirds, see Table 6.

Table 6

Combined results grouped by thirds and split by belief

	Combined thirds (LB N = 154; HB N = 141)					
	Total hit	Total miss	LB hit	LB miss	HB hit	HB miss
A1-5	729	746	374	396	355	350
A6-11	891	879	454	470	437	409
A12-16	768	706	384	385	384	321
Total	2388	2331	1212	1251	1176	1080
P1-5	354	360	184	207	170	153
P6-10	371	374	192	214	179	160
P11-15	380	331	173	180	207	151
Total	1105	1065	549	601	556	464
T1-5	181	177	96	108	85	69
T6-9	142	153	69	95	73	58
T10-14	178	158	71	86	107	72
Total	501	488	236	289	265	199

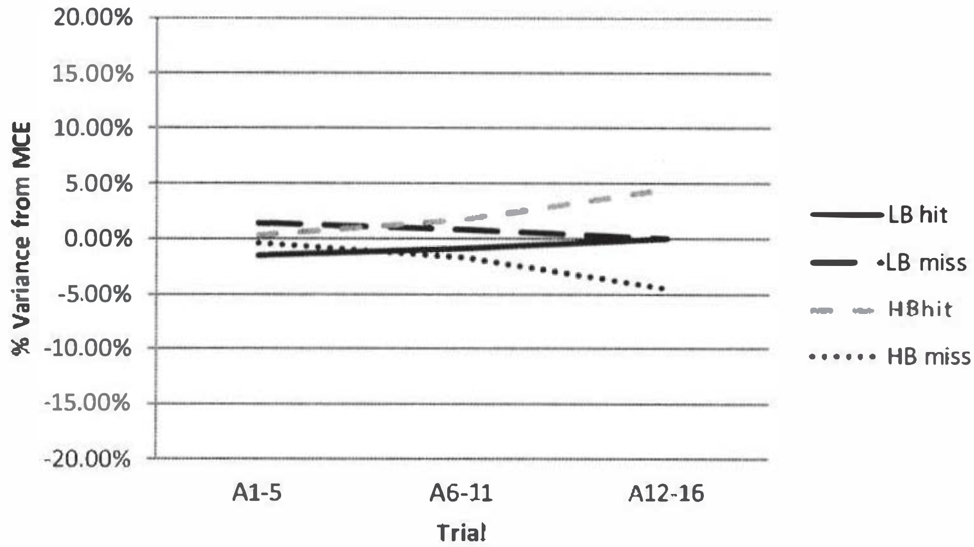

Figure 4. Percent variance from mean chance expectation of combined hits and misses split by belief and grouped by thirds.

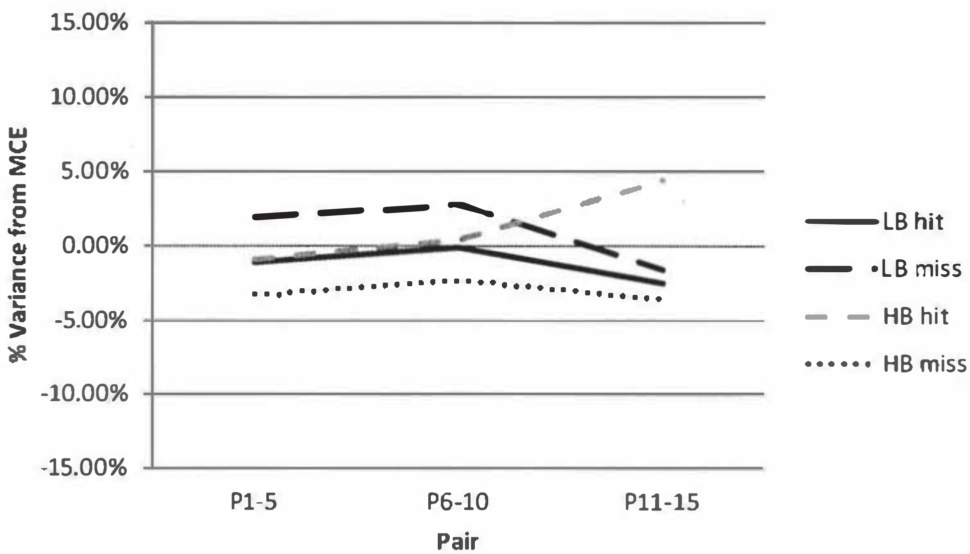

Figure 5. Percent variance from mean chance expectation of combined pairs of hits and misses split by belief and grouped by thirds.

Dataset A: Warner-**Angel**

This dataset consisted of 98 participants total (LB N = 33; HB N = 37) running 16 trials each for a total of 1568 trials. For a complete analysis of the individual trial data, see Warner-Angel (2011). As an examination of strings of hits and misses for the combined datasets showed no significant differences prior to separating participants by belief, the following data analysis examines high and low belief participants separately. Results of the raw trial data split by belief are presented in Table 7 and Figure 6 below.

The total number of pairs of hits for high belief participants was 161, while the total number of pairs of misses was 96. The total number of pairs of hits for low belief participants was 113, while the total number of pairs of misses was 134. A chi-square test indicated a significant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 14.49$ (1, N = 70), $p = .00014$. For totals of pairs of hits and misses, see Table 8 and Figure 7. A post-hoc binomial analysis of 161 pairs of hits (successes) versus 96 pairs of misses for high belief participants resulted in a cumulative probability of $p(X \geq x)$ p = .000030, or three in 100,000.

The total number of triplets of hits for high belief participants was 85, while the total number of triplets of misses was 35. The total number of triplets of hits for low belief participants was 46, while the total number of triplets of misses was 63. A chi-square test indicated a significant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 19.13$ (1, N = 70), $p = .000012$. For totals of triplets of hits and misses, see Table 8. Again, a post-hoc binomial analysis of 85 triplets of hits versus 35 triplets of misses for high belief participants results in a cumulative probability of p = .0000028, or 2.8 in 1 million.

The total number of hits for high belief participants on trials 12 through 16 was 112, while the total number of misses on trials 12 through 16 was 73. The total number of hits for low belief participants on trials 12 through 16 was 82, while the total number of misses on trials 12 through 16 was 83. A chi-square test indicated a significant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 4.1509$ (1, N = 70), $p = .042$. For totals split by belief and grouped into thirds, see Table 9 and Figure 8.

The total number of hits for high belief participants on pairs 11 through 15 was 65, while the total number of misses on pairs 11 through 15 was 24. The total number of hits for low belief participants on pairs 11 through 15 was 41, while the total number of misses on pairs 11 through 15 was 48. A chi-square test indicated a significant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 13.434$ (1, N = 70), $p = .00025$. For totals split by belief and grouped into thirds, see Table 9 and Figure 9.

The total number of hits for high belief participants on triplets 10 through 14 was 36, while the total number of misses on triplets 10 through 14 was 8. The total number of hits for low belief participants on triplets 10 through 14 was 19, while the total number of misses on triplets 10 through 14 was 28. A chi-square test indicated a significant relationship between hit and miss rates when comparing high and low belief participants, χ^2 = 16.2845 (1, N = 70), p = .000055. For totals split by belief and grouped into thirds, see Table 9 and Figure 10. As the analyses presented here were significant, support was found for hypothesis C. Additionally, the presence of a strong order effect in the data was established.

Table 7

Raw trial scores split by belief for Dataset A

	Dataset A (LB N = 33; HB N = 37)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Trial 1	32	38	12 (36.4)	21 (63.6)	20 (54.1)	17 (45.9)
Trial 2	28	42	12 (36.4)	21 (63.6)	16 (43.2)	21 (56.8)
Trial 3	37	33	16 (48.5)	17 (51.5)	21 (56.8)	16 (43.2)
Trial 4	39	31	16 (48.5)	17 (51.5)	23 (62.2)	14 (37.8)
Trial 5	39	31	19 (57.6)	14 (42.4)	20 (54.1)	17 (45.9)
Trial 6	39	31	19 (57.6)	14 (42.4)	20 (54.1)	17 (45.9)
Trial 7	36	34	19 (57.6)	14 (42.4)	17 (45.9)	20 (54.1)
Trial 8	33	37	9 (27.3)	24 (72.7)	24 (64.9)	13 (35.1)
Trial 9	38	32	17 (51.5)	16 (48.5)	21 (56.8)	16 (43.2)
Trial 10	34	36	18 (54.5)	15 (45.5)	16 (43.2)	21 (56.8)
Trial 11	32	38	14 (42.4)	19 (57.6)	18 (48.6)	19 (51.4)
Trial 12	36	34	12 (36.4)	21 (63.6)	24 (64.9)	13 (35.1)
Trial 13	38	32	14 (42.4)	19 (57.6)	24 (64.9)	13 (35.1)

Trial 14	40	30	17 (51.5)	16 (48.5)	23 (62.2)	14 (37.8)
Trial 15	44	26	19 (57.6)	14 (42.4)	25 (67.6)	12 (32.4)
Trial 16	36	34	20 (60.6)	13 (39.4)	16 (43.2)	21 (56.8)
Total	581	539	253 (47.9)	275 (52.1)	328 (55.4)	264 (44.6)

Note: LB = low belief; HB = high belief

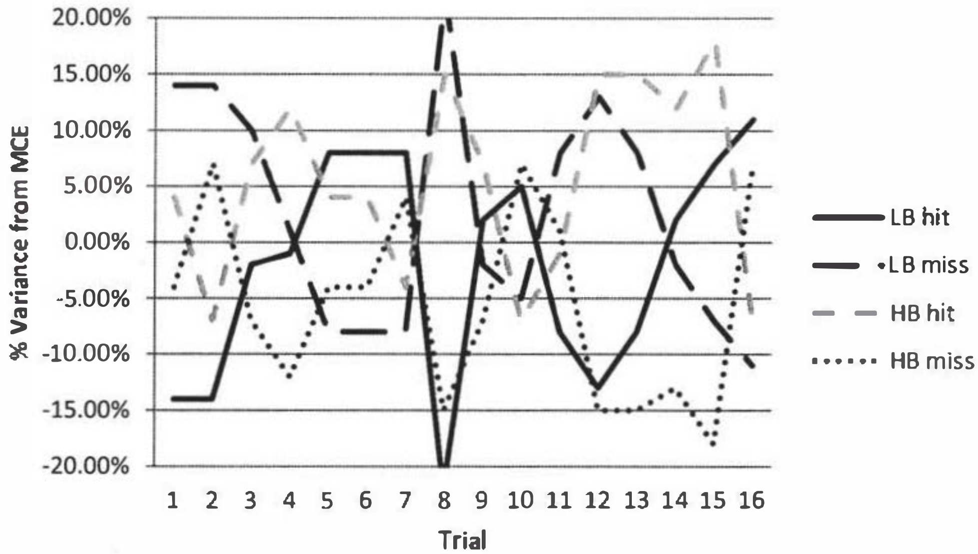

Figure 6. Raw trial scores split by belief for Dataset A.

Table 8

Pairs and triplets split by belief for Dataset A

	Dataset A (LB N = 33; HB N = 37)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Pair 1	10	20	4 (12.1)	13 (39.4)	6 (16.2)	7 (18.9)

Pair 2	11	16	5 (15.2)	10 (30.3)	6 (16.2)	6 (16.2)
Pair 3	19	13	5 (15.2)	6 (18.2)	14 (37.8)	7 (18.9)
Pair 4	20	12	8 (24.2)	6 (18.2)	12 (32.4)	6 (16.2)
Pair 5	22	14	12 (36.4)	7 (21.2)	10 (27.0)	7 (18.9)
Pair 6	22	17	11 (33.3)	6 (18.2)	11 (29.7)	11 (29.7)
Pair 7	15	16	4 (12.1)	9 (27.3)	11 (29.7)	7 (18.9)
Pair 8	18	17	5 (15.2)	12 (36.4)	13 (35.1)	5 (13.5)
Pair 9	17	15	9 (27.3)	7 (21.2)	8 (21.6)	8 (21.6)
Pair 10	14	18	9 (27.3)	10 (30.3)	5 (13.5)	8 (21.6)
Pair 11	16	18	5 (15.2)	12 (36.4)	11 (29.7)	6 (16.2)
Pair 12	19	15	6 (18.2)	13 (39.4)	13 (35.1)	2 (5.4)
Pair 13	23	15	8 (24.2)	10 (30.3)	15 (40.5)	5 (13.5)
Pair 14	26	12	10 (30.3)	7 (21.2)	16 (43.2)	5 (13.5)
Pair 15	22	12	12 (36.4)	6 (18.2)	10 (27.0)	6 (16.2)
Total	274	230	113 (22.8)	134 (27.1)	161 (29.0)	96 (17.3)

	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Triple 1	5	8	2 (6.1)	7 (21.2)	3 (8.1)	1 (2.7)
Triple 2	5	6	0 (0)	4 (12.1)	5 (13.5)	2 (5.4)
Triple 3	9	4	1 (3.0)	1 (3.0)	8 (21.6)	3 (8.1)
Triple 4	11	7	6 (18.2)	4 (12.1)	5 (13.5)	3 (8.1)
Triple 5	12	7	7 (21.2)	2 (6.1)	5 (13.5)	5 (13.5)
Triple 6	8	9	1 (3.0)	4 (12.1)	7 (18.9)	5 (13.5)

Triple 7	11	6	3 (9.1)	4 (12.1)	8 (21.6)	2 (5.4)
Triple 8	6	8	2 (6.1)	5 (15.2)	4 (10.8)	3 (8.1)
Triple 9	9	7	5 (15.2)	4 (12.1)	4 (10.8)	3 (8.1)
Triple 10	6	9	3 (9.1)	6 (18.2)	3 (8.1)	3 (8.1)
Triple 11	9	10	1 (3.0)	9 (27.3)	8 (21.6)	1 (2.7)
Triple 12	11	8	3 (9.1)	7 (21.2)	8 (21.6)	1 (2.7)
Triple 13	15	5	6 (18.2)	3 (9.1)	9 (24.3)	2 (5.4)
Triple 14	14	4	6 (18.2)	3 (9.1)	8 (21.6)	1 (2.7)
Total	131	98	46 (10.0)	63 (13.6)	85 (16.4)	35 (6.8)

Note: LB = low belief; HB = high belief

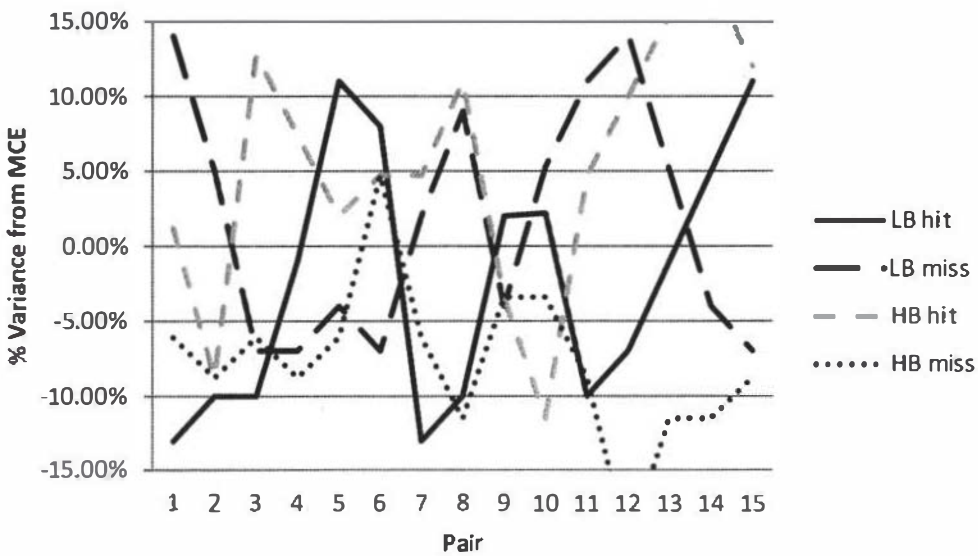

Figure 7. Pairs split by belief for Dataset A.

Table 9

Results split by belief and grouped for Dataset A

Dataset A (LB N = 33; HB N = 37)

	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
A1-5	175	175	75 (45.5)	90 (54.5)	100 (54.1)	85 (45.9)
A6-11	212	208	96 (48.5)	102 (51.5)	116 (52.3)	106 (47.7)
A12-16	194	156	82 (49.7)	83 (50.3)	112 (60.5)	73 (39.5)
Total	581	539	253 (47.9)	275 (52.1)	328 (55.4)	264 (44.6)
P1-5	82	75	34 (20.6)	42 (25.5)	48 (25.9)	33 (17.8)
P6-10	86	83	38 (23.0)	44 (26.7)	48 (25.9)	39 (21.1)
P11-15	106	72	41 (24.8)	48 (29.1)	65 (35.1)	24 (13.0)
Total	274	230	113 (22.8)	134 (27.1)	161 (29.0)	96 (17.3)
T1-5	42	32	16 (10.0)	18 (10.9)	26 (14.1)	14 (7.6)
T6-9	34	30	11 (8.3)	17 (12.9)	23 (15.5)	13 (8.8)
T10-14	55	36	19 (11.5)	28 (17.0)	36 (19.5)	8 (4.3)
Total	131	98	46 (10.0)	63 (13.6)	85 (16.4)	35 (6.8)

Note: LB = low belief; HB = high belief; A1-5 = Trials 1 through 5; A6-11 = Trials 6 through 11; A12-16 = Trials 12 through 16; P1-5 = Pairs 1 through 5; P6-10 = Pairs 6 through 10; P11-15 = Pairs 11 through 15; T1-5 = Triplets 1 through 5; T6-9 = Triplets 6 through 9; T10-14 = Triplets 10 through 14

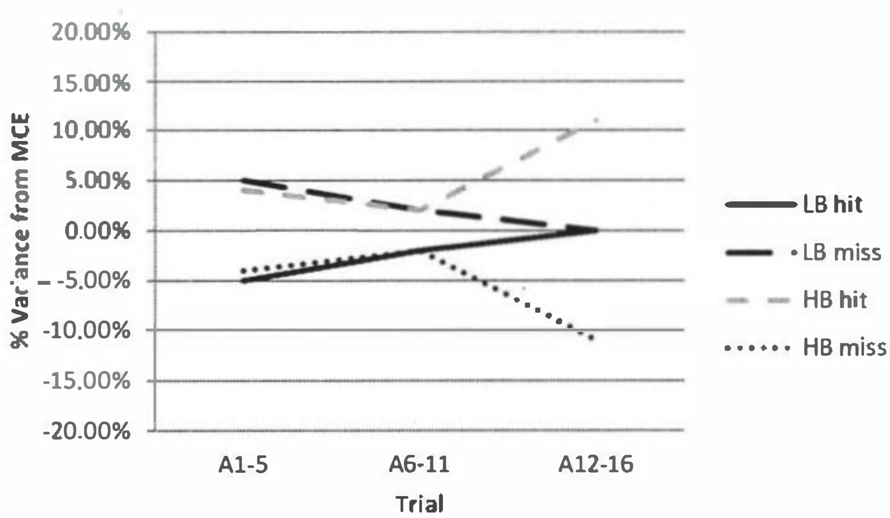

Figure 8. Trials split by belief and grouped for Dataset A.

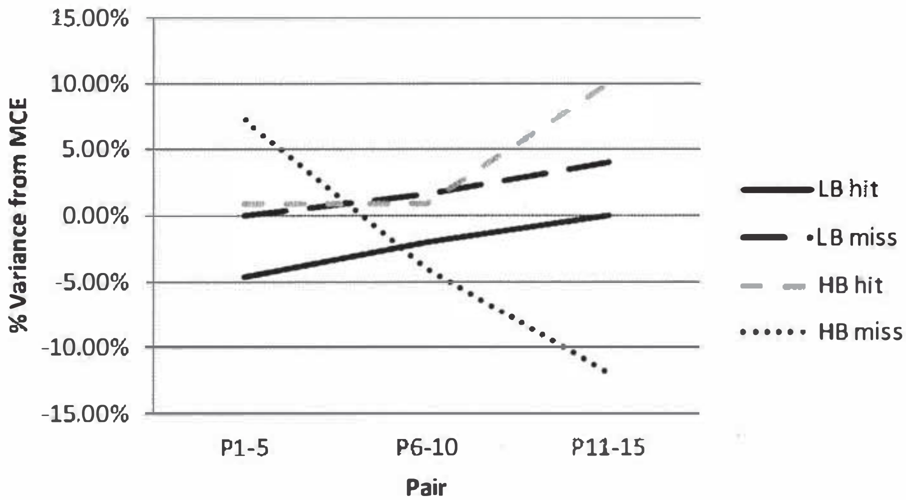

Figure 9. Pairs split by belief and grouped for Dataset A.

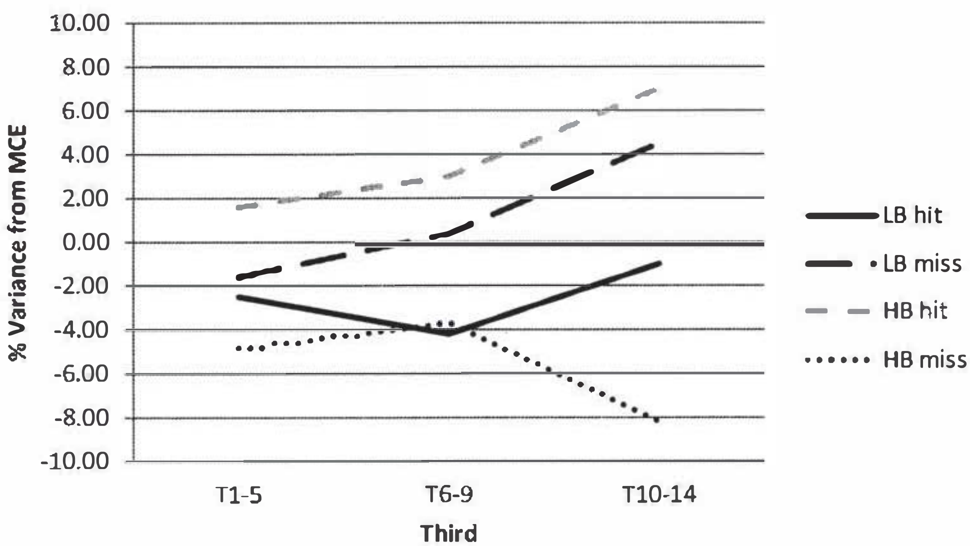

Figure 10. Triplets split by belief and grouped for Dataset A.

Dataset B: Hinman

This dataset involved 53 participants total (LB N = 23; HB N = 25) running 16 trials each for a total of 848 trials. Overall, participants scored 418 hits (for a hit rate of

49.3%). For analysis of individual trials, see Hinman (2017). Results of the raw trial data split by belief are also presented in Table 10 and Figure 11 below.

The total number of pairs of hits for high belief participants was 96, while the total number of pairs of misses was 84. The total number of pairs of hits for low belief participants was 86, while the total number of pairs of misses was 95. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 1.2227$ (1, N = 48), $p = .2688$; however, the results were in the expected direction. For totals of pairs of hits and misses, see Table 11 and Figure 12.

The total number of triplets of hits for high belief participants was 49, while the total number of triplets of misses was 38. The total number of triplets of hits for low belief participants was 41, while the total number of triplets of misses was 50. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 2.2587$ (1, N = 48), $p = .1329$; however, the results were in the expected direction. For totals of triplets of hits and misses, see Table 11.

The total number of hits for high belief participants on trials 12 through 16 was 66, while the total number of misses on trials 12 through 16 was 59. The total number of hits for low belief participants on trials 12 through 16 was 60, while the total number of misses on trials 12 through 16 was 55. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 0.094$ (1, N = 48), $p = .9227$. For totals split by belief and grouped into thirds, see Table 12 and Figure 13.

The total number of hits for high belief participants on pairs 11 through 15 was 35, while the total number of misses on pairs 11 through 15 was 30. The total number of hits for low belief participants on pairs 11 through 15 was 33, while the total number of misses on pairs 11 through 15 was 20. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 0.8472$ (1, N = 48), $p = .3573$. For totals split by belief and grouped into thirds, see Table 12 and Figure 14.

The total number of hits for high belief participants on triplets 10 through 14 was 19, while the total number of misses on triplets 10 through 14 was 17. The total number of hits for low belief participants on triplets 10 through 14 was 19, while the total number of misses on triplets 10 through 14 was 8. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 1.995$ (1, N = 48), $p = .1578$. For totals split by belief and grouped into thirds, see Table 12. As none of the analyses presented here were significant, support was not found for hypothesis C.

Table 10

Raw trial scores split by belief for Dataset B

	Dataset B (LB N = 23; HB N = 25)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Trial 1	23	25	13 (56.5)	10 (43.5)	10 (40.0)	15 (60.0)
Trial 2	20	28	11 (47.8)	12 (52.2)	9 (36.0)	16 (64.0)
Trial 3	22	26	12 (52.2)	11 (47.8)	10 (40.0)	15 (60.0)
Trial 4	23	25	8 (34.8)	15 (65.2)	15 (60.0)	10 (40.0)
Trial 5	24	24	8 (34.8)	15 (65.2)	16 (64.0)	9 (36.0)
Trial 6	27	21	9 (39.1)	14 (60.9)	18 (72.0)	7 (28.0)
Trial 7	25	23	11 (47.8)	12 (52.2)	14 (56.0)	11 (44.0)
Trial 8	25	23	10 (43.5)	13 (56.5)	15 (60.0)	10 (40.0)
Trial 9	16	32	5 (21.7)	18 (78.3)	11 (44.0)	14 (56.0)
Trial 10	23	25	15 (65.2)	8 (34.8)	8 (32.0)	17 (68.0)
Trial 11	32	16	17 (73.9)	6 (26.1)	15 (60.0)	10 (40.0)
Trial 12	24	24	13 (56.5)	10 (43.5)	11 (44.0)	14 (56.0)
Trial 13	22	26	13 (56.5)	10 (43.5)	9 (36.0)	16 (64.0)
Trial 14	27	21	13 (56.5)	10 (43.5)	14 (56.0)	11 (44.0)
Trial 15	27	21	12 (52.2)	11 (47.8)	15 (60.0)	10 (40.0)
Trial 16	26	22	9 (39.1)	14 (60.9)	17 (68.0)	8 (32.0)
Total	386	382	179 (48.6)	189 (51.4)	207 (51.8)	193 (48.2)

Note: LB = low belief; HB = high belief

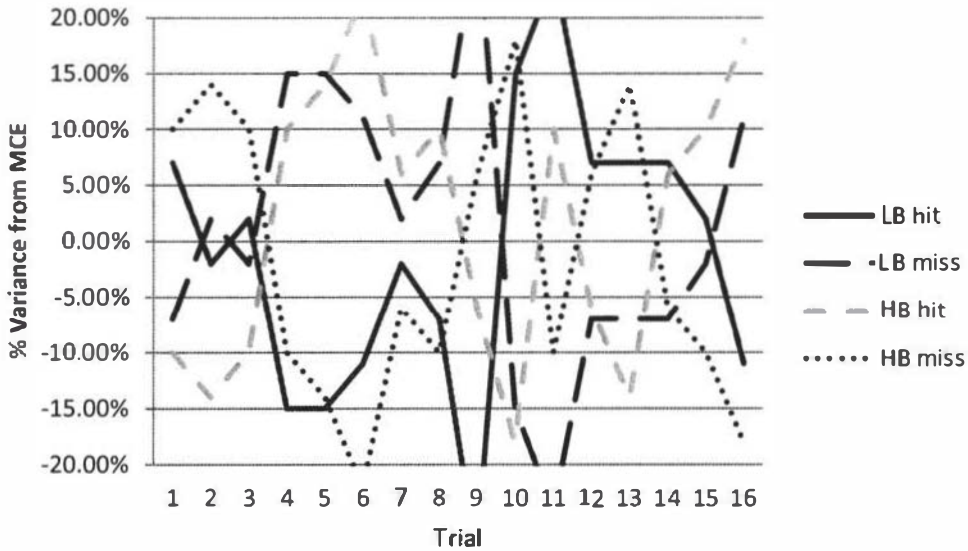

Figure 11. Raw trial scores split by belief for Dataset B.

Table 11

Pairs and triplets split by belief for Dataset B

	Dataset B (LB N = 23; HB N = 25)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Pair 1	9	14	7 (30.4)	6 (26.1)	2 (8.0)	8 (32.0)
Pair 2	9	15	6 (26.1)	6 (26.1)	3 (12.0)	9 (36.0)
Pair 3	11	14	5 (21.7)	8 (34.8)	6 (24.0)	6 (24.0)
Pair 4	12	13	4 (17.4)	11 (47.8)	8 (32.0)	2 (8.0)
Pair 5	14	11	3 (13.0)	9 (39.1)	11 (44.0)	2 (8.0)
Pair 6	12	8	4 (17.4)	7 (30.4)	8 (32.0)	1 (4.0)

Pair 7	15	13	7 (30.4)	9 (39.1)	8 (32.0)	4 (16.0)
Pair 8	9	16	3 (13.0)	11 (47.8)	6 (24.0)	5 (20.0)
Pair 9	7	16	3 (13.0)	6 (26.1)	4 (16.0)	10 (40.0)
Pair 10	16	9	11 (47.8)	2 (8.7)	5 (20.0)	7 (28.0)
Pair 11	16	8	9 (39.1)	2 (8.7)	7 (28.0)	6 (24.0)
Pair 12	11	13	7 (30.4)	4 (17.4)	4 (16.0)	9 (36.0)
Pair 13	11	10	6 (26.1)	3 (13.0)	5 (20.0)	7 (28.0)
Pair 14	16	10	7 (30.4)	5 (21.7)	9 (36.0)	5 (20.0)
Pair 15	14	9	4 (17.4)	6 (26.1)	10 (40.0)	3 (12.0)
Total	182	179	86 (24.9)	95 (27.5)	96 (25.6)	84 (22.4)

	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Triple 1	5	8	4 (17.4)	4 (17.4)	1 (4.0)	4 (16.0)
Triple 2	3	8	2 (8.7)	4 (17.4)	1 (4.0)	4 (16.0)
Triple 3	8	8	3 (13.0)	7 (30.4)	5 (20.0)	1 (4.0)
Triple 4	6	8	1 (4.3)	7 (30.4)	5 (20.0)	1 (4.0)
Triple 5	7	4	3 (13.0)	4 (17.4)	4 (16.0)	0 (0)
Triple 6	10	4	4 (17.4)	4 (17.4)	6 (24.0)	0 (0)
Triple 7	5	10	2 (8.7)	8 (34.8)	3 (12.0)	2 (8.0)
Triple 8	3	6	1 (4.3)	2 (8.7)	2 (8.0)	4 (16.0)
Triple 9	5	7	2 (8.7)	2 (8.7)	3 (12.0)	5 (20.0)
Triple 10	10	4	6 (26.1)	1 (4.3)	4 (16.0)	3 (12.0)
Triple 11	8	5	6 (26.1)	0 (0)	2 (8.0)	5 (20.0)

Triple 12	7	6	3 (13.0)	1 (4.3)	4 (16.0)	5 (20.0)
Triple 13	5	5	2 (8.7)	2 (8.7)	3 (12.0)	3 (12.0)
Triple 14	8	5	2 (8.7)	4 (17.4)	6 (24.0)	1 (4.0)
Total	90	88	41 (12.7)	50 (15.5)	49 (14.0)	38 (10.9)

Note: LB = low belief; HB = high belief

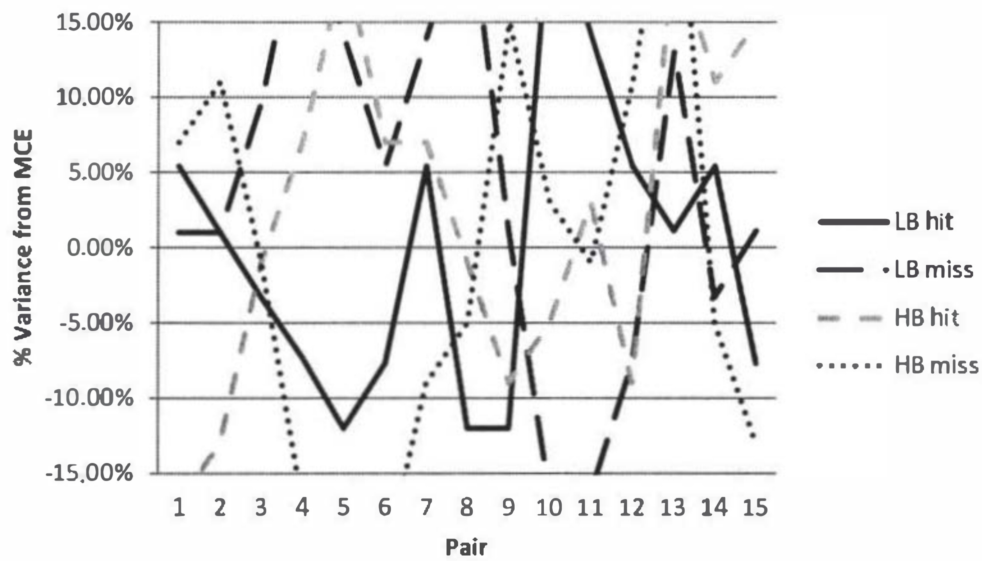

Figure 12. Pair scores split by belief for Dataset B.

Table 12

Results split by belief and grouped for Dataset B

	Dataset B (LB N = 23; HB N = 25)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
A1-5	112	128	52 (45.2)	63 (54.8)	60 (48.0)	65 (52.0)
A6-11	148	140	67 (48.6)	71 (51.4)	81 (54.0)	69 (46.0)
A12-16	126	114	60 (52.2)	55 (47.8)	66 (52.8)	59 (47.2)
Total	386	382	179 (48.6)	189 (51.4)	207 (51.8)	193 (48.2)
P1-5	55	67	25 (21.7)	40 (34.8)	30 (24.0)	27 (21.6)
P6-10	59	62	28 (24.4)	35 (30.4)	31 (24.8)	27 (21.6)
P11-15	68	50	33 (28.7)	20 (17.4)	35 (28.0)	30 (24.0)
Total	182	179	86 (24.9)	95 (27.5)	96 (25.6)	84 (22.4)
T1-5	29	36	13 (11.3)	26 (22.6)	16 (12.8)	10 (8.0)
T6-9	23	27	9 (9.8)	16 (17.4)	14 (14.0)	11 (11.0)
T10-14	38	25	19 (16.5)	8 (7.0)	19 (15.2)	17 (13.6)
Total	90	88	41 (12.7)	50 (15.5)	49 (14.0)	38 (10.9)

Note: LB = low belief; HB = high belief; A1-5 = Trials 1 through 5; A6-11 = Trials 6 through 11; A12-16 = Trials 12 through 16; P1-5 = Pairs 1 through 5; P6-10 = Pairs 6 through 10; P11-15 = Pairs 11 through 15; T1-5 = Triplets 1 through 5; T6-9 = Triplets 6 through 9; T10-14 = Triplets 10 through 14

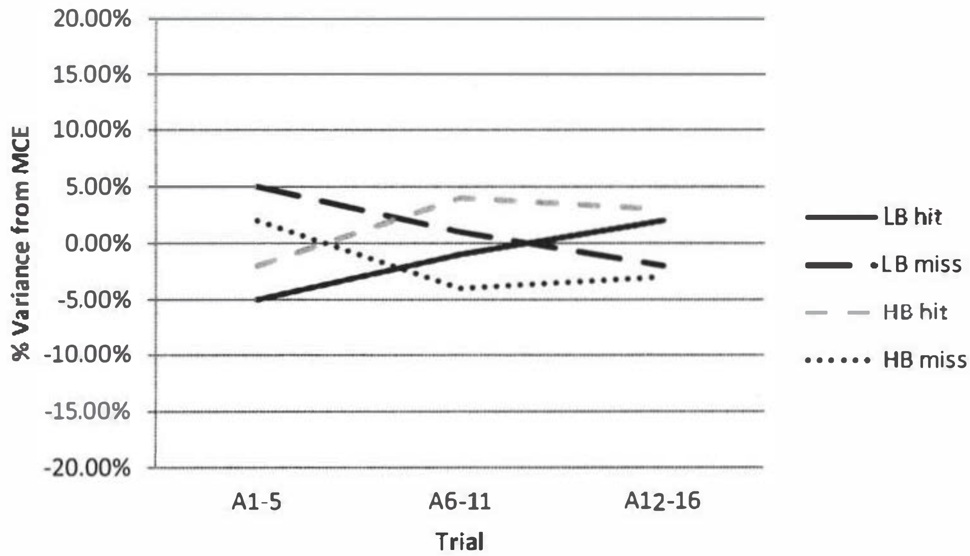

Figure 13. Trials split by belief and grouped for Dataset B.

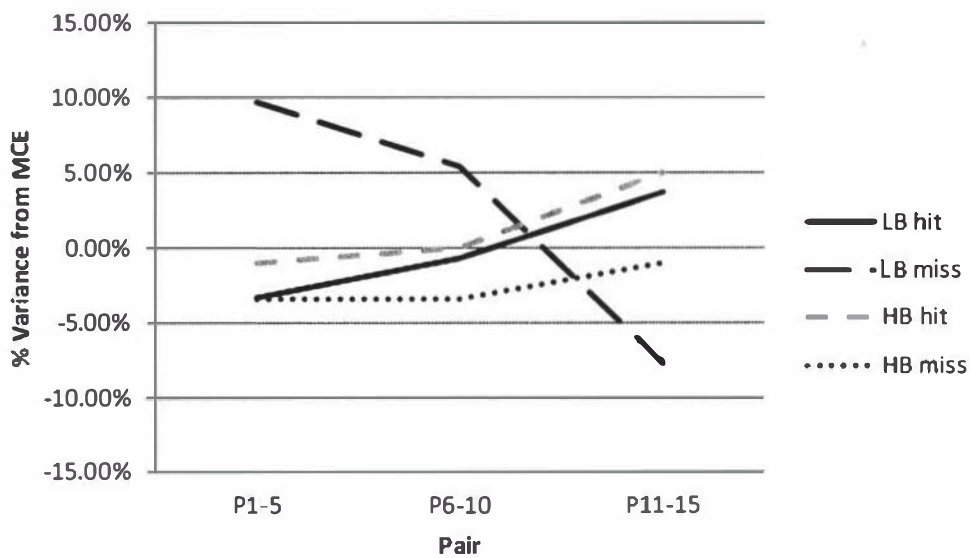

Figure 14. Pairs split by belief and grouped for Dataset B.

Dataset C: Unpublished Data

This dataset involved 90 participants total (LB N = 44; HB N = 30) completing 16 trials each for a total of 1440 trials. Overall, participants scored 732 hits (for a hit rate of 50.8%). The total number of hits for high belief participants was 248, while the total number of misses was 232. The total number of hits for low belief participants was 351, while the total number of misses was 352, which was not significant. For totals of hits and misses, see Table 13 and Figure 15.

The total number of pairs of hits for high belief participants was 117, while the total number of pairs of misses was 106. The total number of pairs of hits for low belief participants was 162, while the total number of pairs of misses was 170. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 0.7192$ (1, N = 74), $p = .3964$; however, the results were in the expected direction. For totals of pairs of hits and misses, see Table 14 and Figure 16.

The total number of triplets of hits for high belief participants was 52, while the total number of triplets of misses was 48. The total number of triplets of hits for low belief participants was 71, while the total number of triplets of misses was 86. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 1.1244$ (1, N = 74), $p = .2890$; however, the results were in the expected direction. For totals of triplets of hits and misses, see Table 14.

The total number of hits for high belief participants on trials 12 through 16 was 84, while the total number of misses on trials 12 through 16 was 66. The total number of hits for low belief participants on trials 12 through 16 was 107, while the total number of misses on trials 12 through 16 was 112. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 1.8184$ (1, N = 74), $p = .1775$; however, the results were in the expected direction. For totals split by belief and grouped into thirds, see Table 15 and Figure 17.

The total number of hits for high belief participants on pairs 11 through 15 was 45, while the total number of misses on pairs 11 through 15 was 30. The total number of hits for low belief participants on pairs 11 through 15 was 44, while the total number of misses on pairs 11 through 15 was 48. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 2.46$ (1, N = 74), $p = .1168$; however, the results were in the expected direction. For totals split by belief and grouped into thirds, see Table 15 and Figure 18.

The total number of hits for high belief participants on triplets 10 through 14 was 25, while the total number of misses on triplets 10 through 14 was 17. The total number of hits for low belief participants on triplets 10 through 14 was 17, while the total number of misses on triplets 10 through 14 was 22. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 2.0565$ (1, N = 74), $p = .1516$; however, the results were in the expected direction. For totals split by belief and grouped into thirds, see Table 15. As none of the data analyzed here were significant, support was not found for hypothesis C.

Table 13

Raw trial scores split by belief for Dataset C

	Dataset C (LB N = 44; HB N = 30)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Trial 1	45	29	27 (61.4)	17 (38.6)	18 (60.0)	12 (40.0)
Trial 2	40	34	24 (54.5)	20 (45.5)	16 (53.3)	14 (46.7)
Trial 3	35	39	22 (50.0)	22 (50.0)	13 (43.3)	17 (56.7)
Trial 4	36	38	25 (56.8)	19 (43.2)	11 (36.7)	19 (63.3)
Trial 5	35	39	17 (38.6)	27 (61.4)	18 (60.0)	12 (40.0)
Trial 6	39	35	24 (54.5)	20 (45.5)	15 (50.0)	15 (50.0)
Trial 7	35	39	20 (45.5)	24 (54.5)	15 (50.0)	15 (50.0)
Trial 8	34	40	18 (40.9)	26 (59.1)	16 (53.3)	14 (46.7)
Trial 9	37	37	21 (47.7)	23 (52.3)	16 (53.3)	14 (46.7)
Trial 10	33	41	21 (47.7)	23 (52.3)	12 (40.0)	18 (60.0)
Trial 11	39	35	25 (56.8)	19 (43.2)	14 (46.7)	16 (53.3)
Trial 12	37	37	20 (45.5)	24 (54.5)	17 (56.7)	13 (43.3)
Trial 13	36	38	22 (50.0)	22 (50.0)	14 (46.7)	16 (53.3)
Trial 14	39	35	23 (52.3)	21 (47.7)	16 (53.3)	14 (46.7)

Trial 15	39	35	18 (40.9)	26 (59.1)	21 (70.0)	9 (30.0)
Trial 16	40	33	24 (54.5)	19 (43.2)	16 (53.3)	14 (46.7)
Total	599	584	351 (49.9)	352 (50.0)	248 (51.7)	232 (48.3)

Note: LB = low belief; HB = high belief

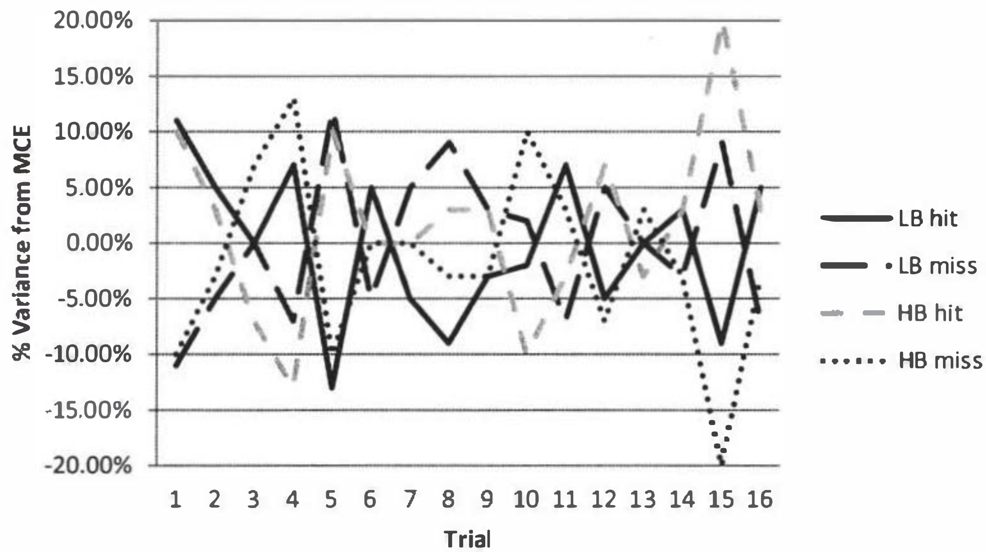

Figure 15. Trial scores split by belief for Dataset C.

Table 14

Pairs and triplets split by belief for Dataset C

	Dataset C (LB N = 44; HB N = 30)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Pair 1	28	17	15 (34.1)	8 (18.2)	13 (43.3)	9 (30.0)
Pair 2	19	18	11 (25.0)	9 (20.5)	8 (26.7)	9 (30.0)

Pair 3	16	19	13 (29.5)	10 (22.7)	3 (10.0)	9 (30.0)
Pair 4	15	18	11 (25.0)	13 (29.5)	4 (13.3)	5 (16.7)
Pair 5	18	18	9 (20.5)	12 (27.3)	9 (30.0)	6 (20.0)
Pair 6	24	24	14 (31.8)	14 (31.8)	10 (33.3)	10 (33.3)
Pair 7	14	19	8 (18.2)	14 (31.8)	6 (20.0)	5 (16.7)
Pair 8	21	24	11 (25.0)	16 (36.4)	10 (33.3)	8 (26.7)
Pair 9	15	19	12 (27.3)	14 (31.8)	3 (10.0)	5 (16.7)
Pair 10	20	22	14 (31.8)	12 (27.3)	6 (20.0)	10 (33.3)
Pair 11	17	15	10 (22.7)	9 (20.5)	7 (23.3)	6 (20.0)
Pair 12	14	16	6 (13.6)	8 (18.2)	9 (30.0)	8 (26.7)
Pair 13	18	17	11 (25.0)	10 (22.7)	7 (23.3)	7 (23.3)
Pair 14	21	17	9 (20.5)	12 (27.3)	12 (40.0)	5 (16.7)
Pair 15	19	13	8 (18.2)	9 (20.5)	11 (36.7)	4 (13.3)
Total	279	276	162 (24.5)	170 (25.8)	117 (26.0)	106 (23.6)

	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Triple 1	11	9	5 (11.4)	3 (6.8)	6 (20.0)	6 (20.0)
Triple 2	7	10	6 (13.6)	5 (11.4)	1 (3.3)	5 (16.7)
Triple 3	8	10	6 (13.6)	9 (20.5)	2 (6.7)	1 (3.3)
Triple 4	9	8	7 (15.9)	6 (13.6)	2 (6.7)	2 (6.7)
Triple 5	10	12	5 (11.4)	7 (15.9)	5 (16.7)	5 (16.7)
Triple 6	10	11	6 (13.6)	8 (18.2)	4 (13.3)	3 (10.0)
Triple 7	9	11	5 (11.4)	9 (20.5)	4 (13.3)	2 (6.7)

Triple 8	8	12	6 (13.6)	9 (20.5)	2 (6.7)	3 (10.0)
Triple 9	9	12	8 (18.2)	8 (18.2)	1 (3.3)	4 (13.3)
Triple 10	6	9	4 (9.1)	6 (13.6)	2 (6.7)	3 (10.0)
Triple 11	9	7	3 (6.8)	3 (6.8)	6 (20.0)	4 (13.3)
Triple 12	8	6	3 (6.8)	3 (6.8)	5 (16.7)	3 (10.0)
Triple 13	9	11	3 (6.8)	7 (15.9)	6 (20.0)	4 (13.3)
Triple 14	10	6	4 (9.1)	3 (6.8)	6 (20.0)	3 (10.0)
Total	123	134	71 (11.5)	86 (14.0)	52 (12.4)	48 (11.4)

Note: LB = low belief; HB = high belief

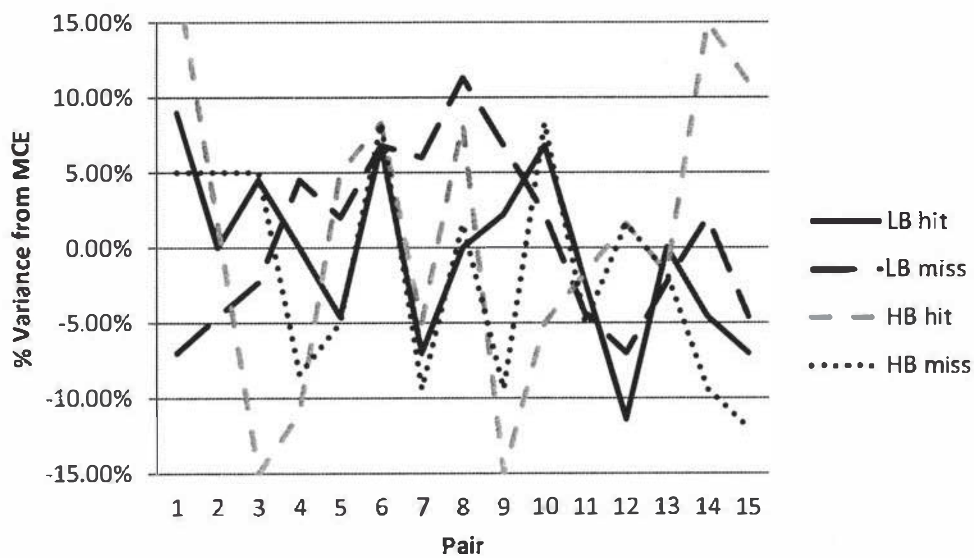

Figure 16. Pairs split by belief for Dataset C.

Table 15

Results split by belief and grouped by thirds for Dataset C

	Dataset C (LB N = 44; HB N = 30)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
A1-5	191	179	115 (52.3)	105 (47.7)	76 (50.7)	74 (49.3)
A6-11	217	227	129 (48.9)	135 (51.1)	88 (48.9)	92 (51.1)
A12-16	191	178	107 (48.6)	112 (50.9)	84 (56.0)	66 (44.0)
Total	599	584	351 (49.9)	352 (50.0)	248 (51.7)	232 (48.3)
P1-5	96	90	59 (26.8)	52 (23.6)	37 (24.7)	38 (25.3)
P6-10	94	108	59 (26.8)	70 (31.8)	35 (23.3)	38 (25.3)
P11-15	89	78	44 (20.0)	48 (21.8)	45 (30.0)	30 (20.0)
Total	279	276	162 (24.5)	170 (25.8)	117 (26.0)	106 (23.6)
T1-5	45	49	29 (13.2)	30 (13.6)	16 (10.7)	19 (12.7)
T6-9	36	46	25 (14.2)	34 (19.3)	11 (9.2)	12 (10.0)
T10-14	42	39	17 (7.7)	22 (10.0)	25 (16.7)	17 (11.3)
Total	123	134	71 (11.5)	86 (14.0)	52 (12.4)	48 (11.4)

Note: LB = low belief; HB = high belief; A1-5 = Trials 1 through 5; A6-11 = Trials 6 through 11; A12-16 = Trials 12 through 16; P1-5 = Pairs 1 through 5; P6-10 = Pairs 6 through 10; P11-15 = Pairs 11 through 15; T1-5 = Triplets 1 through 5; T6-9 = Triplets 6 through 9; T10-14 = Triplets 10 through 14

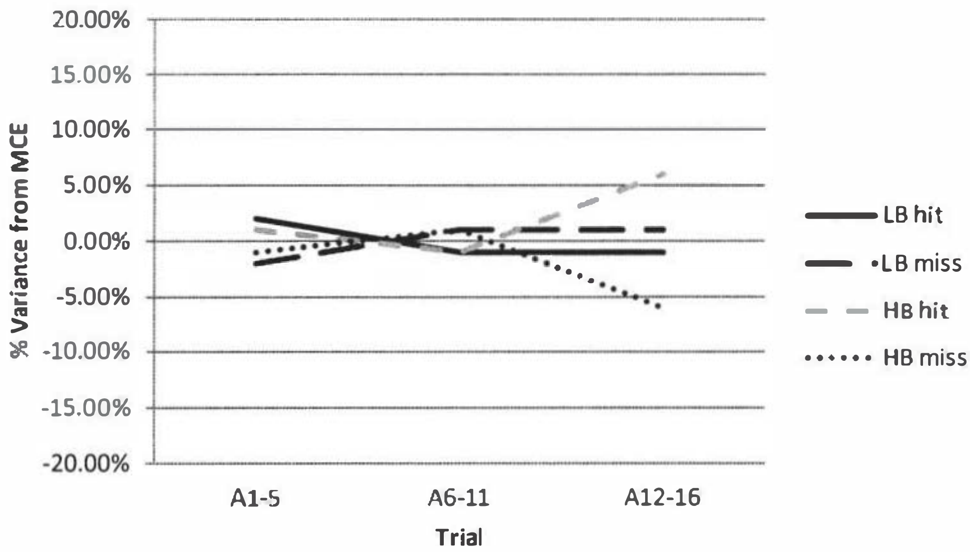

Figure 17. Trials split by belief and grouped by thirds for Dataset C.

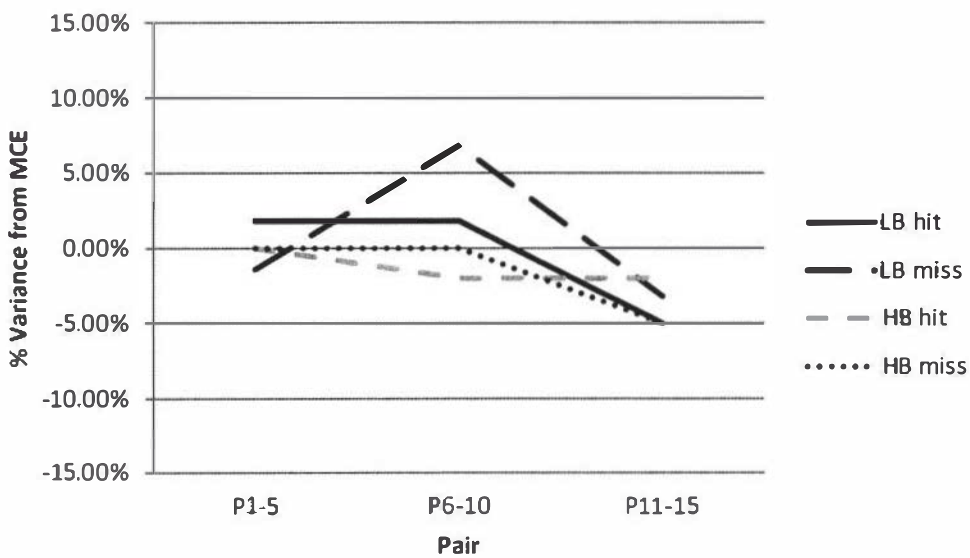

Figure 18. Pairs split by belief and grouped by thirds for Dataset C.

Dataset D: Unpublished Data

This dataset involved 107 participants total (LB N = 54; HB N = 49) running 16 trials each for a total of 1712 trials. Overall, participants scored 858 hits (for a hit rate of 50.1%). The total number of hits for high belief participants was 393, while the total number of misses was 391. The total number of hits for low belief participants was 429, while the total number of misses was 435, which was not significant. For totals of hits and misses, see Table 16 and Figure 19.

The total number of pairs of hits for high belief participants was 182, while the total number of pairs of misses was 178. The total number of pairs of hits for low belief participants was 188, while the total number of pairs of misses was 202. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 0.4137$ (1, N = 103), $p = .5201$. For totals of pairs of hits and misses, see Table 17 and Figure 20.

The total number of triplets of hits for high belief participants was 79, while the total number of triplets of misses was 78. The total number of triplets of hits for low belief participants was 78, while the total number of triplets of misses was 90. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 0.4918$ (1, N = 103), $p = .4831$. For totals of triplets of hits and misses, see Table 17.

The total number of hits for high belief participants on trials 12 through 16 was 122, while the total number of misses on trials 12 through 16 was 123. The total number of hits for low belief participants on trials 12 through 16 was 135, while the total number of misses on trials 12 through 16 was 135. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 0.0021$ (1, N = 103), $p = .9631$. For totals split by belief and grouped into thirds, see Table 18 and Figure 21.

The total number of hits for high belief participants on pairs 11 through 15 was 62, while the total number of misses on pairs 11 through 15 was 67. The total number of hits for low belief participants on pairs 11 through 15 was 55, while the total number of misses on pairs 11 through 15 was 64. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 0.0844$ (1, N = 103), $p = .7714$. For totals split by belief and grouped into thirds, see Table 18 and Figure 22.

The total number of hits for high belief participants on triplets 10 through 14 was 27, while the total number of misses on triplets 10 through 14 was 30. The total number of hits for low belief participants on triplets 10 through 14 was 16, while the total number of misses on triplets 10 through 14 was 28. A chi-square test indicated a nonsignificant relationship between hit and miss rates when comparing high and low belief participants, $\chi^2 = 1.23$ (1, N = 103), $p = .2674$. For totals split by belief and grouped into thirds, see Table 18. As none of the analyses presented here were significant, support was not found for hypothesis C.

Table 16

Raw trial scores split by belief for Dataset D

	Dataset D (LB N = 54; HB N = 49)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Trial 1	50	53	28 (51.9)	26 (48.1)	22 (44.9)	27 (55.1)
Trial 2	48	55	26 (48.1)	28 (51.9)	22 (44.9)	27 (55.1)
Trial 3	49	54	26 (48.1)	28 (51.9)	23 (46.9)	26 (53.1)
Trial 4	55	48	27 (50.0)	27 (50.0)	28 (57.1)	21 (42.9)
Trial 5	49	54	25 (46.3)	29 (53.7)	24 (49.0)	25 (51.0)
Trial 6	56	47	27 (50.0)	27 (50.0)	29 (59.2)	20 (40.8)
Trial 7	53	50	27 (50.0)	27 (50.0)	26 (53.1)	23 (46.9)
Trial 8	54	49	25 (46.3)	29 (53.7)	29 (59.2)	20 (40.8)
Trial 9	53	50	28 (51.9)	26 (48.1)	25 (51.0)	24 (49.0)
Trial 10	52	51	30 (55.6)	24 (44.4)	22 (44.9)	27 (55.1)
Trial 11	46	57	25 (46.3)	29 (53.7)	21 (42.9)	28 (57.1)
Trial 12	49	54	22 (40.7)	32 (59.3)	27 (55.1)	22 (44.9)
Trial 13	50	53	24 (44.4)	30 (55.6)	26 (53.1)	23 (46.9)
Trial 14	50	53	26 (48.1)	28 (51.9)	24 (49.0)	25 (51.0)
Trial 15	49	54	29 (53.7)	25 (46.3)	20 (40.8)	29 (59.2)
Trial 16	59	44	34 (63.0)	20 (37.0)	25 (51.0)	24 (49.0)
Total	822	826	429 (49.7)	435 (50.3)	393 (50.1)	391 (49.9)

Note: LB = low belief; HB = high belief

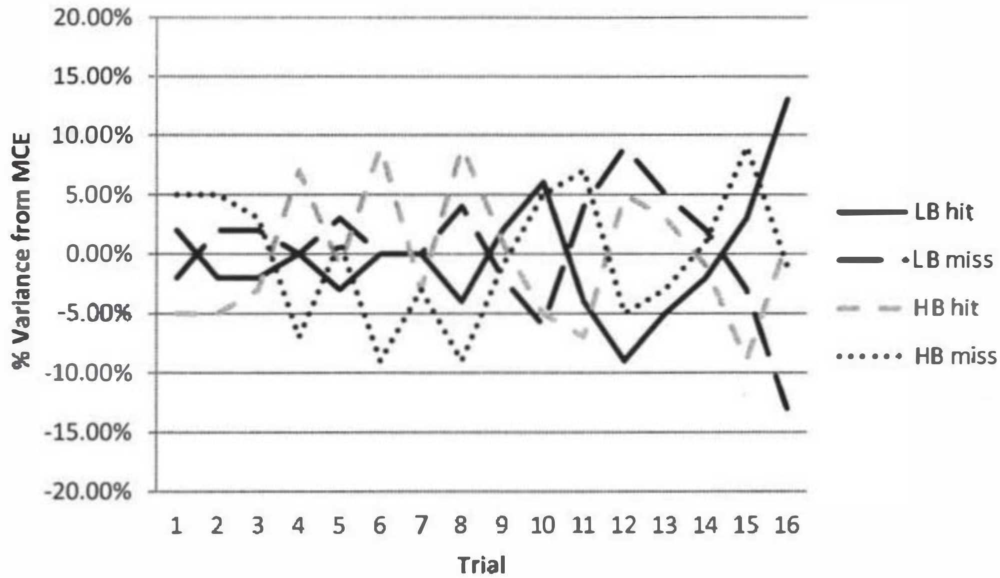

Figure 19. Trial scores split by belief for Dataset D.

Table 17

Pairs and triplets split by belief for Dataset D

	Dataset D (LB N = 54; HB N = 49)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Pair 1	25	30	16 (29.6)	16 (29.6)	9 (18.4)	14 (28.6)
Pair 2	20	26	12 (22.2)	14 (25.9)	8 (16.3)	12 (24.5)
Pair 3	29	28	15 (27.8)	16 (29.6)	14 (28.6)	12 (24.5)
Pair 4	25	24	13 (24.1)	15 (27.8)	12 (24.5)	9 (18.4)
Pair 5	22	20	10 (18.5)	12 (22.2)	12 (24.5)	8 (16.3)
Pair 6	29	23	12 (22.2)	12 (22.2)	17 (34.7)	11 (22.4)

Pair 7	26	22	11 (20.4)	13 (24.1)	15 (30.6)	9 (18.4)
Pair8	27	23	13 (24.1)	14 (25.9)	14 (28.6)	9 (18.4)
Pair9	28	26	16 (29.6)	12 (22.2)	12 (24.5)	14 (28.6)
Pair 10	22	27	15 (27.8)	14 (25.9)	7 (14.3)	13 (26.5)
Pair 11	19	27	7 (13.0)	14 (25.9)	12 (24.5)	13 (26.5)
Pair 12	25	29	10 (18.5)	18 (33.3)	15 (30.6)	11 (22.4)
Pair 13	24	27	11 (20.4)	15 (27.8)	13 (26.5)	12 (24.5)
Pair 14	21	25	10 (18.5)	9 (16.7)	11 (22.4)	16 (32.7)
Pair 15	28	23	17 (31.5)	8 (14.8)	11 (22.4)	15 (30.6)
Total	370	380	188 (23.2)	202 (24.9)	182 (24.8)	178 (24.2)

	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
Triple 1	14	14	11 (20.4)	8 (14.8)	3 (6.1)	6 (12.2)
Triple 2	14	12	8 (14.8)	7 (13.0)	6 (12.2)	5 (10.2)
Triple 3	14	17	9 (16.7)	10 (18.5)	5 (10.2)	7 (14.3)
Triple 4	12	7	5 (9.3)	4 (7.4)	7 (14.3)	3 (6.1)
Triple 5	11	10	5 (9.3)	5 (9.3)	6 (12.2)	5 (10.2)
Triple 6	13	10	4 (7.4)	6 (11.1)	9 (18.4)	4 (8.2)
Triple 7	10	14	4 (7.4)	9 (16.7)	6 (12.2)	5 (10.2)
Triple 8	15	11	8 (14.8)	6 (11.1)	7 (14.3)	5 (10.2)
Triple 9	11	15	8 (14.8)	7 (13.0)	3 (6.1)	8 (16.3)
Triple 10	7	13	2 (3.7)	6 (11.1)	5 (10.2)	7 (14.3)
Triple 11	9	12	2 (3.7)	6 (11.1)	7 (14.3)	6 (12.2)

Triple 12	10	13	4 (7.4)	9 (16.7)	6 (12.2)	4 (8.2)
Triple 13	7	11	3 (5.6)	5 (9.3)	4 (8.2)	6 (12.2)
Triple 14	10	9	5 (9.3)	2 (3.7)	5 (10.2)	7 (14.3)
Total	157	168	78 (10.3)	90 (11.9)	79 (11.5)	78 (11.4)

Note: LB = low belief; HB = high belief

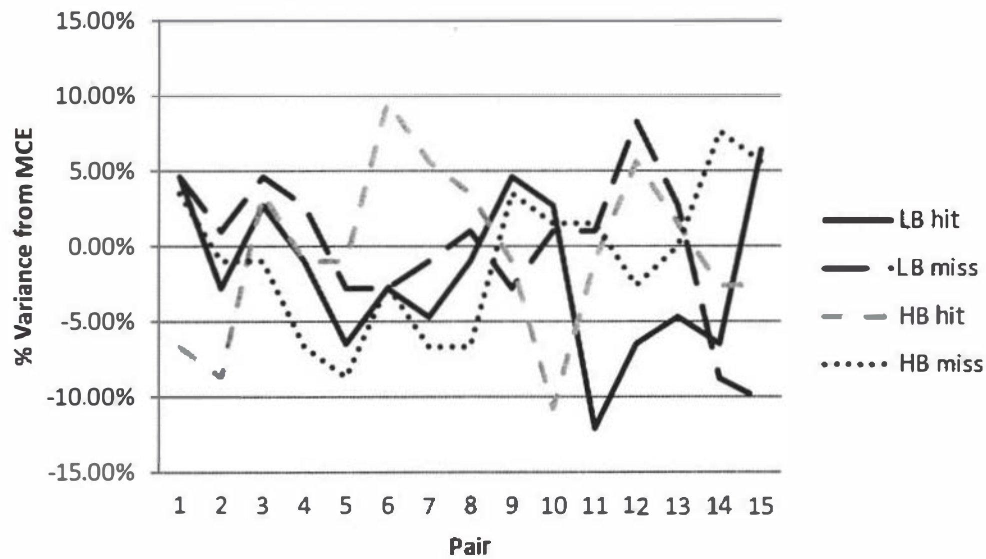

Figure 20. Pairs split by belief for Dataset D.

Table 18

Results split by belief and grouped for Dataset D

	Dataset D (LB N = 54; HB N = 49)					
	Total hits	Total miss	LB hit (%)	LB miss (%)	HB hit (%)	HB miss (%)
A1-5	251	264	132 (48.9)	138 (51.1)	119 (48.6)	126 (51.4)
A6-11	314	304	162 (50.0)	162 (50.0)	152 (51.7)	142 (48.3)
A12-16	257	258	135 (50.0)	135 (50.0)	122 (49.8)	123 (50.2)
Total	822	826	429 (49.7)	435 (50.3)	393 (50.1)	391 (49.9)
P1-5	121	128	66 (24.4)	73 (27.0)	55 (22.5)	55 (22.5)
P6-10	132	121	67 (24.8)	65 (24.1)	65 (26.5)	56 (22.9)
P11-15	117	131	55 (20.4)	64 (23.7)	62 (25.3)	67 (27.3)
Total	370	380	188 (23.2)	202 (24.9)	182 (24.8)	178 (24.2)
T1-5	65	60	38 (14.1)	34 (12.6)	27 (11.0)	26 (10.6)
T6-9	49	50	24 (11.1)	28 (13.0)	25 (12.8)	22 (11.2)
T10-14	43	58	16 (5.9)	28 (10.4)	27 (11.0)	30 (12.3)
Total	157	168	78 (10.3)	90 (11.9)	79 (11.5)	78 (11.4)

Note: LB = low belief; HB = high belief; A1-5 = Trials 1 through 5; A6-11 = Trials 6 through 11; A12-16 = Trials 12 through 16; P1-5 = Pairs 1 through 5; P6-10 = Pairs 6 through 10; P11-15 = Pairs 11 through 15; T1-5 = Triplets 1 through 5; T6-9 = Triplets 6 through 9; T10-14 = Triplets 10 through 14

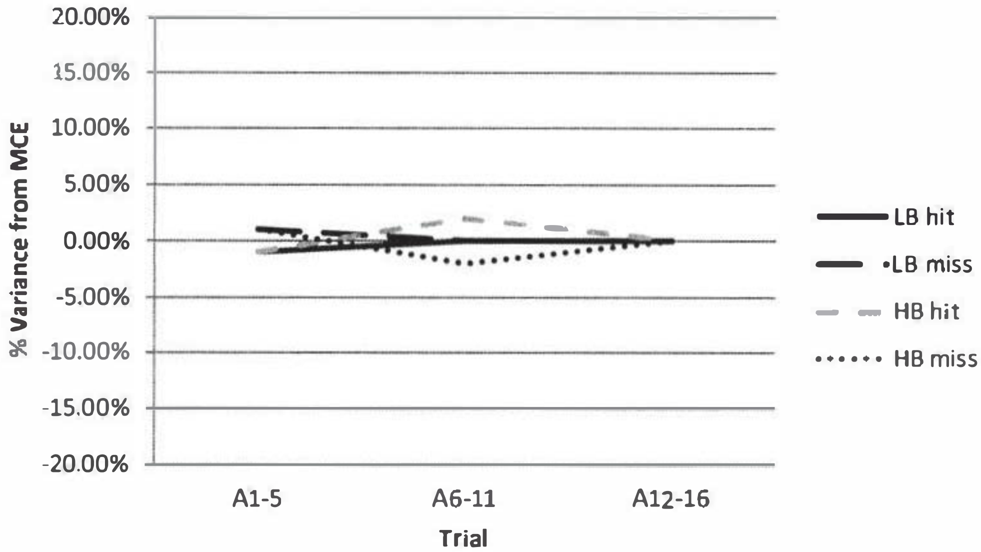

Figure 21. Trial scores split by belief and grouped by thirds for Dataset D.

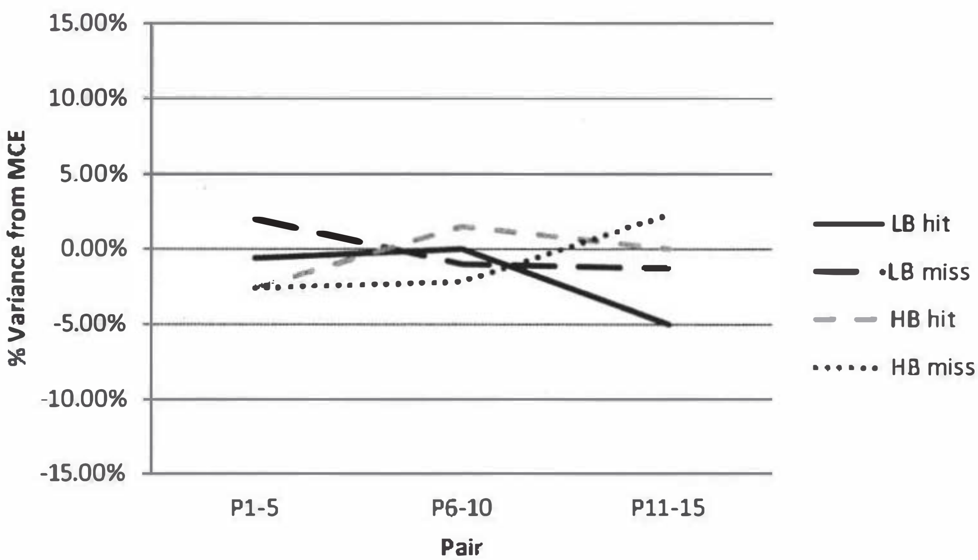

Figure 22. Pair scores split by belief and grouped by thirds for Dataset D.

Discussion

The current study is a reanalysis of four semesters of data collected from the Eastern Illinois University (EIU) parapsychology lab. The goals of the current study were to determine whether there were more strings of hits and misses than expected by chance in a psi task (A); whether there was a difference between the number of strings of hits and the number of strings of misses (B); and whether participant scores on a 7-point belief in psi scale would be related to the frequency of hits and misses (C). Each participant in the experiments that were analyzed completed 16 experimental trials, resulting in 15 possible pairs (strings of 2) of hits or misses and 14 possible triplets (strings of 3) of hits or misses.

When the combined datasets were analyzed using binomial probability, they did not vary significantly from mean chance expectation (MCE). We had expected an MCE of a 50% hit rate ($M = 8.00$) for trials and found ($M = 8.08$) for hits and ($M = 7.92$) for misses. For pairs, we expected an MCE of 25% ($M = 3.75$) and found ($M = 3.76$) for hits and ($M = 3.65$) for misses. For triplets, we expected an MCE of 12.5% ($M = 1.75$) and found ($M = 1.73$) for hits and ($M = 1.68$) for misses. As these results are nonsignificant, my first two hypotheses were not supported.

Although the results of analyses for all four datasets combined appeared random at first, splitting the datasets by belief provided very strong support for my third hypothesis and the existence of nonrandom occurrences in the datasets. A t test comparing hits of high and low belief participants for individual trial data was significant at $p = .039$, showing that high belief participants scored more hits than low belief participants. Chi-squares were used to analyze pairs and triplets split by belief because these data were not interval or ratio data. Chi-squares comparing hits and misses between high and low belief participants for the pairs and triplets were significant at $p = .0016$ and $p = .00014$ respectively. This shows strong support for the tendency of hits to follow hits and misses to follow misses, and suggests that the effect of the split by belief actually increased with longer strings. Post-hoc binomial analysis showed that differences between pairs and triplets of hits and misses for high belief participants were significant at p = .000030 and p = .0000028 respectively.

In order to further examine the effect of split by belief, strings of four and five were also examined. Quads (strings of four hits or misses) and quintuplets (strings of five hits or misses) showed similar albeit weaker results in the same direction as the pairs and triplets analyses. Chi-square analyses comparing hits and misses between high and low belief participants revealed that quads just barely failed to reach significance at $p = .055$, while quintuplets were significant at $p = .019$. It is unclear why the significance level dropped for quads and increased again for quintuplets, but because these results are less significant than the triplets, it seems likely that there are diminishing effects regarding the length of extended strings. Analysis of further extended strings in the datasets was hampered by small sample sizes, as strings of six or more proved to be relatively rare. This suggests a need for a participant pool far larger than a combined pool of 348 participants when examining extended strings of hits and misses, in order to achieve a usable sample size for statistical analyses.

Schmeidler & Murphy (1945) found similar results to those found in the current study when considering participant belief and performance on an extrasensory perception (ESP) task, with ESP believers scoring more hits over time and ESP disbelievers scoring more misses over time. It is interesting to note that 70 years later, the current study found results similar to a much older study that used less rigorous experimental methods. Luke, Delanoy, and Sherwood (2008) found that belief in psi was significantly and positively correlated with precognition task performance. Bunfill (2006) found that belief in psi influenced scoring decline rates, with high belief participants showing slower declines than low belief participants, which may help to explain the emergence of significant results when combining the datasets split by belief. The current study also replicates previous findings from the lab that show significant evidence of high belief participants hitting more often than low belief participants (Warner-Angel, 2016; McWhorter, 2005). Taken together, these findings suggest that high belief participants tend to hit more often on psi tasks than low belief participants.

Visual inspection of the graphic representation of the datasets split by belief suggested significant variance from MCE toward the end of runs. This may have been due to participants becoming acclimated to the mental telepathy task and thus producing greater variance near the end of their runs. As the trial and pair data showed considerable variability when they were graphed, grouping the experimental runs into thirds was useful in establishing the presence of a position effect. The individual trials, pairs, and triplets were grouped together into a beginning, middle, and end of a given experimental run. When chi-square tests comparing hits and misses between high and low belief participants were conducted, the last third of the individual trials was not significant at $p = .082$; however, the last third of the pairs and triplets were significant at $p = .018$ and $p = .0076$ respectively. The effect that was observed was not due to low belief participants' scoring rates, as they appeared to converge on zero variance from MCE in the last third. However, high belief participants scored considerably more hits and fewer misses during the last third of their runs, suggesting that they had improved over time. It is also noteworthy that a position effect was found for strings of hits and misses but not for the individual trials.

Parker (2006) also found an order effect for psi trials, although the order effect in his data was with respect to the order of participants in an experimental set rather than the order of trials in each run. He had hypothesized an acclimation and fatigue effect across the entire dataset, and found that the first and last participants in a set did significantly worse than other participants, providing support for his hypothesis. While I did not consider the order of participants in each dataset, an analysis of hit and miss rates considered in the context of participant order might be useful to include in future iterations of this type of analysis.

In addition to an analysis of the combined datasets, analyses were conducted on each of the four datasets separately. As with the combined datasets, chi-square tests were conducted for pairs and triplets split by belief as well as for the last third of trials, pairs, and triplets. Overall, high belief participants in Dataset A scored more hits than misses, and their low belief counterparts scored more misses than hits. The presence of a position effect was established for Dataset A, and the size of the effect increased for pairs and increased again for triplets. While the other datasets that were analyzed failed to reach significance, Datasets B and C had results that were in the same direction as those that were found for Dataset A, showing a trend in scoring that accounted for the significant results found in the combined analyses. However, the results for Dataset D appeared to have no directionality, with high and low belief participants scoring roughly equal numbers of hits and misses. It is important to mention that there was considerable variability across the datasets that were analyzed. This may have been due to the fact that participants who were enrolled in the lead researcher's introductory psychology class, had been exposed to information on psi phenomena, and also knew the lead researcher produced significant findings, while participants who were not in his class did not.

Implications for Psi Research

The analyses that were carried out have significant implications for psi research in general and research on psi belief in particular. The data were analyzed in several different ways, but it is noteworthy that conclusive results supporting the existence of psi were not found until the datasets were split by belief. This suggests that categorizing participants by belief may be an essential step when researchers become stymied by apparently elusive psi effects that seem hard to replicate. It is also possible that standard cause-and-effect explanations of psi may not apply at all, which Gruber highlighted by observing that "It seems that we may be looking at some kind of acausal or alternative causality influence on randomness." (Gruber, personal communication, Jan. 6, 2018)

Extended strings of hits and misses may also be related to the psi-missing displacement effect (PMDE), which refers to participants who score significantly below MCE but score above MCE when adjacent targets are considered. This may help explain why analyses of the datasets split into thirds were significant for extended strings but not for individual trials, as well as potentially explaining why the extended string analyses were significant in the first place. Crandall and Hite (1983) found that PMDE occurred when participants were trying to guess target locations on a precognition task, and concluded that PMDE arises from improperly focused psi for those participants who psi-miss. Crandall (1985) conducted further research on PMDE and found that participants who were tested by a passionate experimenter (rather than an indifferent one) showed a greater displacement effect and a larger number of psi-missers who scored above the mean chance expectation (MCE) for displacements. Those who have worked with him would certainly say that Gruber is a passionate experimenter, so it may well be that his lab participants are manifesting significant PMDE effects that may help explain the results that have been found here.

The current study's findings regarding strings of hits and misses may be related to the concepts of the gambler's fallacy and the hot hand fallacy. Recall that the gambler's fallacy states that a different outcome is more likely after a string of identical outcomes, and the hot hand fallacy states that those who "hit" are inherently more likely to continue to do so. They are considered fallacies because, based on probability theory, the chance of a given outcome remains constant. Sundali and Croson (2006) found that people who believe in one of these fallacies typically believe in the other as well. These fallacies appear to be perpetuated by gamblers' perceptions that they do in fact experience streaks of good luck and bad luck. The results of the current study suggest that a gambler's belief in lucky and unlucky streaks may in fact be based on the actual occurrence of strings of hits and misses that deviate from randomness. In addition, current data show that nonrandom strings of hits and misses happen in nearly equal proportions and therefore cancel each other out. Therefore, gamblers may in fact accurately perceive strings of good luck and bad luck, which may promote the desire to gamble, but the gamblers do not end up winning money because these strings of good luck and bad luck appear in equal proportions. Further, it was found that participants with low belief in psi were far more likely to miss three times in a row than to hit three times in a row (chi-square significant at $p = .00014$), suggesting that they are more likely to lose than to win.

After dividing participants based on belief in psi, strings of hits and misses appeared to vary considerably from chance expectation. Because dividing participants by belief led to highly significant findings, future analyses should take into consideration whether or not participants believe in psi. Gruber has suggested that future iterations of this type of analysis could also benefit from splitting the datasets by participant gender (Gruber, personal communication, 2017), which is a factor that I did not consider in my analyses. Past investigations of the effects of gender on psi scoring have revealed that female participants tended to show stronger psi effects than male participants (Conner, 2009; Warner-Angel, 2011). Also, conducting tests of linearity or curvilinearity would help to clarify the directionality and significance of research findings. This reanalysis, along with the previously reported results from the original studies, provides strong support for the existence of mental telepathy or a related psi phenomenon.

Appendix A

Table 19

Raw trial scores

Dataset A (N = 98), Dataset B (N = 53), Dataset C (N = 90), and Dataset D (N = 107)

	Dataset A	Dataset B	Dataset C	Dataset D
Trial 1	44 (44.9)	24 (45.3)	52 (57.8)	53 (49.5)
Trial 2	41 (41.8)	23 (43.4)	49 (54.4)	50 (46.7)
Trial 3	51 (52.0)	25 (47.2)	47 (52.2)	51 (47.7)
Trial 4	56 (57.1)	23 (43.4)	44 (48.9)	57 (53.3)
Trial 5	54 (55.1)	26 (49.1)	45 (50.0)	52 (48.6)
Trial 6	54 (55.1)	30 (56.6)	49 (54.4)	58 (54.2)
Trial 7	50 (51.0)	26 (49.1)	41 (45.6)	54 (50.5)
Trial 8	48 (49.0)	26 (49.1)	41 (45.6)	57 (53.3)
Trial 9	54 (55.1)	18 (34.0)	47 (52.2)	56 (52.3)
Trial 10	47 (48.0)	23 (43.4)	41 (45.6)	54 (50.5)
Trial 11	46 (46.9)	34 (64.2)	46 (51.1)	48 (44.9)
Trial 12	48 (49.0)	26 (49.1)	43 (47.8)	51 (47.7)
Trial 13	53 (54.1)	26 (49.1)	46 (51.1)	52 (48.6)
Trial 14	53 (54.1)	32 (60.4)	45 (50.0)	52 (48.6)
Trial 15	54 (55.1)	27 (50.9)	46 (51.1)	51 (47.7)

Trial 16	51 (52.0)	29 (54.7)	50 (55.6)	62 (57.9)
Total	804 (51.3)	418 (48.7)	732 (50.8)	858 (50.1)

Note: Percentages are displayed next to frequencies in parentheses

Table 20

Pairs and triplets for Dataset A

	Dataset A (N = 98)		
	2 misses (%)	Hit/miss (%)	2 hits (%)
Pair 1	27 (27.6)	57 (58.2)	14 (14.3)
Pair 2	23 (23.5)	58 (59.2)	17 (17.3)
Pair 3	19 (19.4)	51 (52.0)	28 (28.6)
Pair 4	17 (17.3)	52 (53.1)	29 (29.6)
Pair 5	19 (19.4)	50 (51.0)	29 (29.6)
Pair 6	22 (22.4)	48 (49.0)	28 (28.6)
Pair 7	25 (25.5)	48 (49.0)	25 (25.5)
Pair 8	22 (22.4)	50 (51.0)	26 (26.5)
Pair 9	21 (21.4)	53 (54.1)	24 (24.5)
Pair 10	26 (26.5)	51 (52.0)	21 (21.4)
Pair 11	28 (28.6)	46 (46.9)	24 (24.5)
Pair 12	22 (22.4)	51 (52.0)	25 (25.5)
Pair 13	22 (22.4)	46 (46.9)	30 (30.6)

Pair 14	24 (24.5)	41 (41.8)	33 (33.7)
Pair 15	20 (20.4)	51 (52.0)	27 (27.6)

	3 misses(%)	3 hits (%)
Triple 1	12 (12.2)	8 (8.2)
Triple 2	9 (9.2)	8 (8.2)
Triple 3	7 (7.1)	15 (15.3)
Triple 4	9 (9.2)	16 (16.3)
Triple 5	8 (8.2)	15 (15.3)
Triple 6	13 (13.3)	13 (13.3)
Triple 7	10 (10.2)	15 (15.3)
Triple 8	10 (10.2)	11 (11.2)
Triple 9	12 (12.2)	13 (13.3)
Triple 10	15 (15.3)	9 (9.2)
Triple 11	13 (13.3)	12 (12.2)
Triple 12	11 (11.2)	13 (13.3)
Triple 13	11 (11.2)	17 (17.3)
Triple 14	11 (11.2)	18 (18.4)

Table 21

Pairs and triplets for Dataset B

	Dataset B (N = 53)		
	2 misses (%)	Hit/miss (%)	2 hits (%)
Pair 1	16 (30.2)	27 (50.9)	10 (18.9)
Pair 2	16 (30.2)	26 (49.1)	11 (20.8)
Pair 3	16 (30.2)	26 (49.1)	11 (20.8)
Pair 4	16 (30.2)	25 (47.2)	12 (22.6)
Pair 5	12 (22.6)	26 (49.1)	15 (28.3)
Pair 6	10 (18.9)	30 (56.6)	13 (24.5)
Pair 7	16 (30.2)	22 (41.5)	15 (28.3)
Pair 8	18 (34.0)	26 (49.1)	9 (17.0)
Pair 9	19 (35.8)	27 (50.9)	7 (13.2)
Pair 10	12 (22.6)	25 (47.2)	16 (30.2)
Pair 11	9 (17.0)	28 (52.8)	16 (30.2)
Pair 12	14 (26.4)	26 (49.1)	13 (24.5)
Pair 13	10 (18.9)	28 (52.8)	15 (28.3)
Pair 14	10 (18.9)	27 (50.9)	16 (30.2)
Pair 15	11 (20.8)	28 (52.8)	14 (26.4)

	3 misses (%)	3 hits (%)
Triple 1	9 (17.0)	6 (11.3)
Triple 2	9 (17.0)	3 (5.7)
Triple 3	10 (18.9)	8 (15.1)
Triple 4	9 (17.0)	6 (11.3)
Triple 5	5 (9.4)	8 (15.1)
Triple 6	6 (11.3)	10 (18.9)
Triple 7	12 (22.6)	5 (9.4)
Triple 8	8 (15.1)	3 (5.7)
Triple 9	9 (17.0)	5 (9.4)
Triple 10	5 (9.4)	10 (18.9)
Triple 11	5 (9.4)	8 (15.1)
Triple 12	6 (11.3)	9 (17.0)
Triple 13	5 (9.4)	5 (9.4)
Triple 14	5 (9.4)	8 (15.1)

Table 22

Pairs and triplets for Dataset C

	Dataset C (N = 90)		
	2 misses (%)	Hit/miss (%)	2 hits (%)
Pair 1	22 (24.4)	35 (38.9)	33 (36.7)
Pair 2	20 (22.2)	44 (48.9)	26 (28.9)

Pair 3	21 (23.3)	47 (52.2)	22 (24.4)
Pair 4	22 (24.4)	47 (52.2)	21 (23.3)
Pair 5	19 (21.1)	48 (53.3)	23 (25.6)
Pair 6	28 (31.1)	34 (37.8)	28 (31.1)
Pair 7	25 (27.8)	48 (53.3)	17 (18.9)
Pair 8	29 (32.2)	34 (37.8)	27 (30.0)
Pair 9	22 (24.4)	48 (53.3)	20 (22.2)
Pair 10	25 (27.8)	43 (47.8)	22 (24.4)
Pair 11	20 (22.2)	51 (56.7)	19 (21.1)
Pair 12	19 (21.1)	53 (58.9)	18 (20.0)
Pair 13	21 (23.3)	47 (52.2)	22 (24.4)
Pair 14	25 (27.8)	39 (43.3)	26 (28.9)
Pair 15	18 (20.0)	46 (51.1)	25 (27.8)

	0 hits (%)	3 hits (%)
Triple 1	10 (11.1)	15 (16.7)
Triple 2	11 (12.2)	10 (11.1)
Triple 3	11 (12.2)	13 (14.4)
Triple 4	9 (10.0)	11 (12.2)
Triple 5	13 (14.4)	13 (14.4)
Triple 6	13 (14.4)	13 (14.4)
Triple 7	15 (16.7)	12 (13.3)

Triple 8	15 (16.7)	12 (13.3)
Triple 9	13 (14.4)	10 (11.1)
Triple 10	10 (11.1)	6 (6.7)
Triple 11	7 (7.8)	10 (11.1)
Triple 12	8 (8.9)	10 (11.1)
Triple 13	14 (15.6)	12 (13.3)
Triple 14	10 (11.1)	14 (15.6)

Table 23

Pairs and triplets for Dataset D

	Dataset D (N = 107)		
	2 misses (%)	Hit/miss (%)	2 hits (%)
Pair 1	30 (28.0)	51 (47.7)	26 (24.3)
Pair 2	27 (25.2)	59 (55.1)	21 (19.6)
Pair 3	29 (27.1)	48 (44.9)	30 (28.0)
Pair 4	24 (22.4)	57 (53.3)	26 (24.3)
Pair 5	21 (19.6)	62 (57.9)	24 (22.4)
Pair 6	24 (22.4)	54 (50.5)	29 (27.1)
Pair 7	22 (20.6)	59 (55.1)	26 (24.3)
Pair 8	24 (22.4)	53 (49.5)	30 (28.0)
Pair 9	26 (24.3)	52 (48.6)	29 (27.1)
Pair 10	28 (26.2)	56 (52.3)	23 (21.5)

Pair 11	28 (26.2)	59 (55.1)	20 (18.7)
Pair 12	30 (28.0)	51 (47.7)	26 (24.3)
Pair 13	29 (27.1)	52 (48.6)	26 (24.3)
Pair 14	26 (24.3)	59 (55.1)	22 (20.6)
Pair 15	24 (22.4)	53 (49.5)	30 (28.0)

	3 misses (%)	3 hits (%)
Triple 1	14 (13.1)	14 (13.1)
Triple 2	12 (11.2)	14 (13.1)
Triple 3	17 (15.9)	15 (14.0)
Triple 4	7 (6.5)	13 (12.1)
Triple 5	10 (9.3)	11 (10.3)
Triple 6	10 (9.3)	13 (12.1)
Triple 7	14 (13.1)	10 (9.3)
Triple 8	11 (10.3)	16 (15.0)
Triple 9	15 (14.0)	11 (10.3)
Triple 10	14 (13.1)	8 (7.5)
Triple 11	12 (11.2)	9 (8.4)
Triple 12	14 (13.1)	11 (10.3)
Triple 13	12 (11.2)	8 (7.5)
Triple 14	9 (8.4)	11 (10.3)

Table 24

Extended strings

Dataset A (N = 98), Dataset B (N = 53), Dataset C (N = 90), and Dataset D (N = 107)

	Dataset A	Dataset B	Dataset C	Dataset D
4hit1	37 (37.8)	15 (28.3)	31 (34.4)	37 (34.6)
4hit2	4 (4.1)	3 (5.7)	3 (3.3)	3 (2.8)
4miss1	31 (31.6)	16 (30.2)	33 (36.7)	33 (30.8)
4miss2	4 (4.1)	3 (5.7)	5 (5.6)	3 (2.8)
5hit1	17 (17.3)	11 (20.8)	16 (17.8)	21 (19.6)
5hit2	1 (1.0)	1 (1.9)	1 (1.1)	0 (0)
5miss1	11 (11.2)	10 (18.9)	19 (21.1)	12 (11.2)
5miss2	1 (1.0)	1 (1.9)	0 (0)	3 (2.8)
6hit	9 (9.2)	8 (15.1)	6 (6.7)	6 (5.6)
6miss	7 (7.1)	5 (9.4)	8 (8.9)	5 (4.7)
7hit	2 (2.0)	2 (3.8)	5 (5.6)	0 (0)
7miss	4 (4.1)	3 (5.7)	4 (4.4)	3 (2.8)
8hit	2 (2.0)	1 (1.9)	3 (3.3)	0 (0)
8miss	0 (0)	3 (5.7)	2 (2.2)	2 (1.9)
9hit	2 (2.0)	0 (0)	2 (2.2)	0 (0)
9miss	0 (0)	3 (5.7)	1 (1.1)	2 (1.9)
10hit	1 (1.0)	0 (0)	1 (1.1)	0 (0)

10miss	0 (0)	0 (0)	0 (0)	1 (0.9)

Note: Percentages are displayed next to frequencies in parentheses

Table 25

Extended strings split by belief

Dataset A (LB N = 33; HB N = 37), Dataset B (LB N = 23; HB N = 25), Dataset C (LB N = 44; HB N = 30), and Dataset D (LB N = 54; HB N =49)

Dataset A	LB (%)	HB (%)	Dataset B	LB (%)	HB (%)
4hit1	12 (36.4)	17 (45.9)	4hit1	5 (21.7)	10 (40.0)
4hit2	1 (3.0)	2 (5.4)	4hit2	3 (13.0)	0 (0)
4miss1	14 (42.4)	6 (16.2)	4miss1	8 (34.8)	6 (24.0)
4miss2	2 (6.1)	1 (2.7)	4miss2	2 (8.7)	1 (4.0)
5hit1	4 (12.1)	9 (24.3)	5hit1	4 (17.4)	7 (28.0)
5hit2	0 (0)	0 (0)	5hit2	1 (4.3)	0 (0)
5miss1	6 (18.2)	1 (2.7)	5miss1	6 (26.1)	3 (12.0)
5miss2	0 (0)	1 (2.7)	5miss2	1 (4.3)	0 (0)
6hit	1 (3.0)	5 (13.5)	6hit	4 (17.4)	4 (16.0)
6miss	2 (6.1)	2 (5.4)	6miss	3 (13.0)	1 (4.0)
7hit	0 (0)	1 (2.7)	7hit	0 (0)	2 (8.0)
7miss	1 (3.0)	1 (2.7)	7miss	1 (4.3)	1 (4.0)
8hit	0 (0)	1 (2.7)	8hit	0 (0)	1 (4.0)
8miss	0 (0)	0 (0)	8miss	1 (4.3)	1 (4.0)
9hit	0 (0)	1 (2.7)	9hit	0 (0)	0 (0)

9miss	0 (0)	0 (0)	9miss	1 (4.3)	1 (4.0)
10hit	0 (0)	0 (0)	10hit	0 (0)	0 (0)
10miss	0 (0)	0 (0)	10miss	0 (0)	0 (0)
Dataset C	LB (%)	HB (%)	Dataset D	LB(%)	HB (%)
4hit1	14 (31.8)	11 (36.7)	4hit1	16 (29.6)	19 (38.8)
4hit2	0 (0)	1 (3.3)	4hit2	3 (5.6)	0 (0)
4miss1	15 (34.1)	13 (43.3)	4miss1	18 (33.3)	14 (28.6)
4miss2	4 (9.1)	1 (3.3)	4miss2	2 (3.7)	1 (2.0)
5hit1	8 (18.2)	3 (10.0)	5hit1	11 (20.4)	10 (20.4)
5hit2	0 (0)	1 (3.3)	5hit2	0 (0)	0 (0)
5miss1	11 (25.0)	4 (13.3)	5miss1	8 (14.8)	4 (8.2)
5miss2	0 (0)	0 (0)	5miss2	2 (3.7)	1 (2.0)
6hit	3 (6.8)	1 (3.3)	6hit	2 (3.7)	4 (8.2)
6miss	5 (11.4)	1 (3.3)	6miss	3 (5.6)	2 (4.1)
7hit	3 (6.8)	1 (3.3)	7hit	0 (0)	0 (0)
7miss	4 (9.1)	0 (0)	7miss	1 (1.9)	2 (4.1)
8hit	2 (4.5)	0 (0)	8hit	0 (0)	0 (0)
8miss	2 (4.5)	0 (0)	8miss	1 (1.9)	1 (2.0)
9hit	1 (2.3)	0 (0)	9hit	0 (0)	0 (0)
9miss	1 (2.3)	0 (0)	9miss	1 (1.9)	1 (2.0)
10hit	1 (2.3)	0 (0)	10hit	0 (0)	0 (0)
10miss	0 (0)	0 (0)	10miss	0 (0)	1 (2.0)

Note: LB = low belief; HB = high belief

Table 26

Grouped results for all datasets, and combined results for all datasets

Dataset A (LB N = 33; HB N = 37), Dataset B (LB N = 23; HB N = 25), Dataset C (LB N = 44; HB N = 30), and Dataset D (LB N = 54; HB N =49)

Dataset A	Hit	Miss	Dataset B	Hit	Miss	Dataset C	Hit	Miss
A1-5	246	244	A1-5	121	144	A1-5	237	213
A6-11	299	289	A6-11	157	161	A6-11	265	275
A12-16	259	231	A12-16	140	125	A12-16	230	219
P1-5	117	105	P1-5	59	76	P1-5	125	104
P6-10	124	116	P6-10	60	75	P6-10	114	129
P11-15	139	116	P11-15	74	54	P11-15	110	103
T1-5	62	45	T1-5	31	42	T1-5	62	54
T6-9	52	45	T6-9	23	35	T6-9	47	56
T10-14	69	61	T10-14	40	26	T10-14	52	49
Dataset D	Hit	Miss	Combined	Hit	Miss	Hit%	Miss%	
A1-5	263	272	A1-5	867	873	0.49828	0.50172	
A6-11	327	315	A6-11	1048	1040	0.50192	0.49808	
A12-16	268	267	A12-16	897	842	0.51552	0.48391	
P1-5	127	131	P1-5	428	416	0.24598	0.23908	
P6-10	137	124	P6-10	435	444	0.25	0.25517	
P11-15	124	137	P11-15	447	410	0.25690	0.23563	

T1-5	67	60	T1-5	222	201	0.12759	0.11552
T6-9	50	50	T6-9	172	186	0.12356	0.13362
T10-14	47	61	T10-14	208	197	0.11954	0.11322

Appendix B

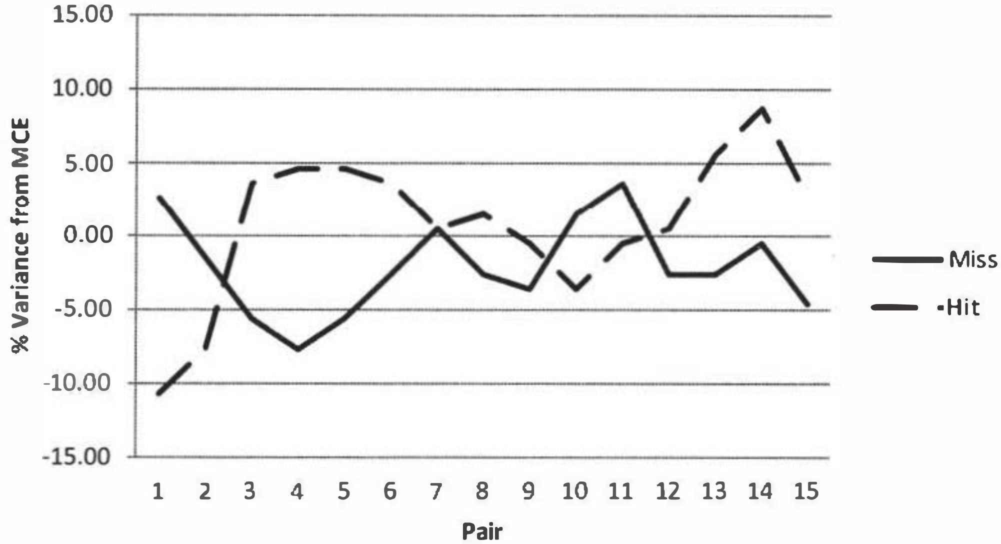

Figure 23. Pairs of hits and misses for Dataset A.

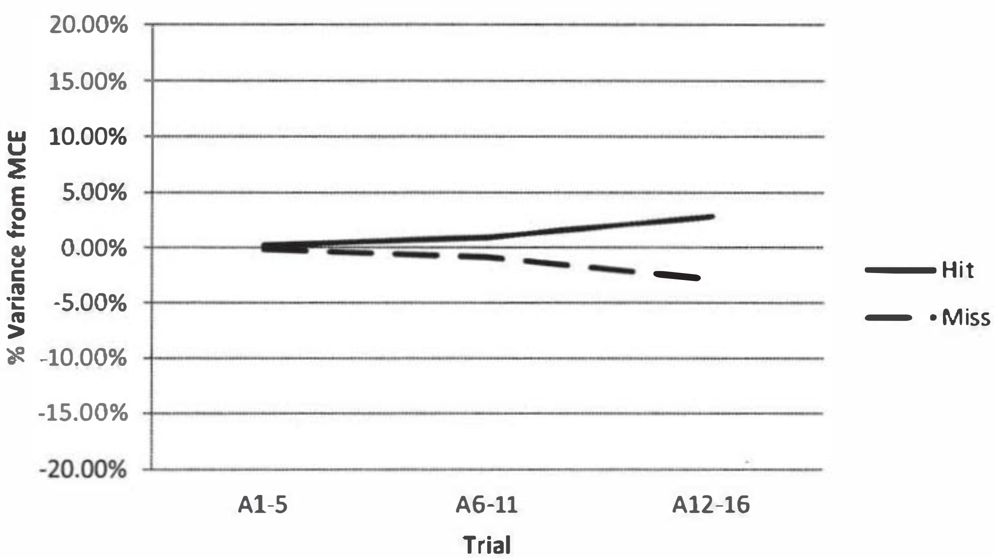

Figure 24. Trials grouped by thirds for Dataset A.

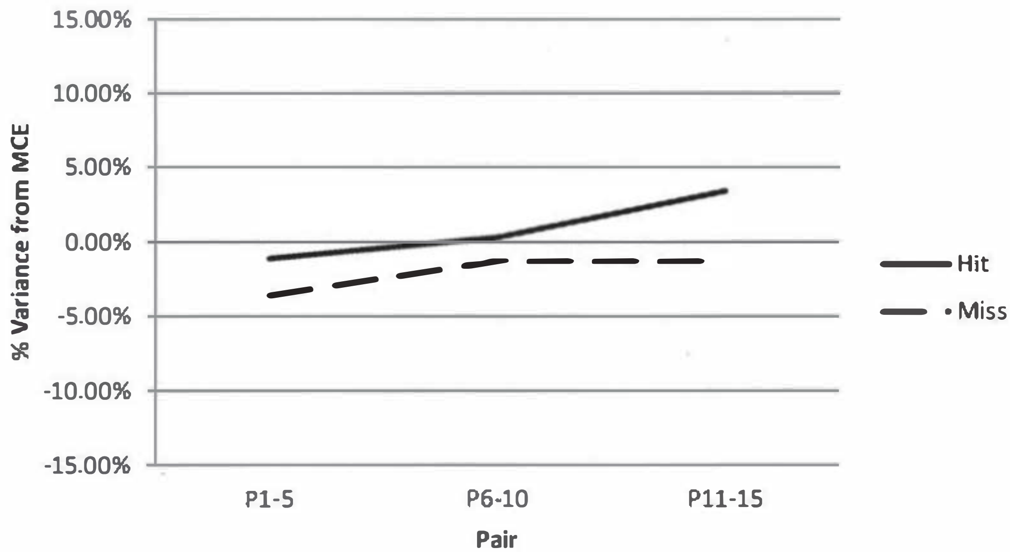

Figure 25. Pairs grouped by thirds for Dataset A.

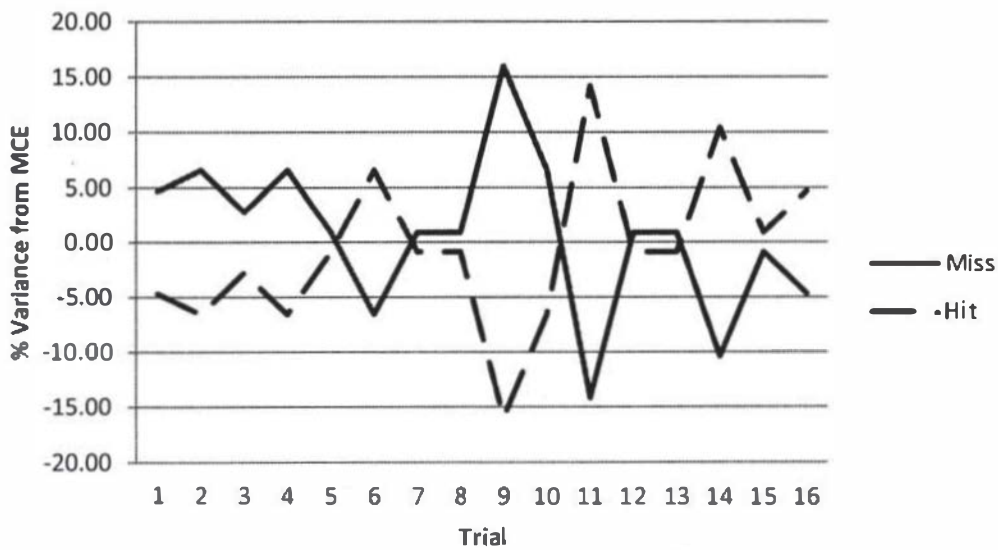

Figure 26. Trial scores for Dataset B.

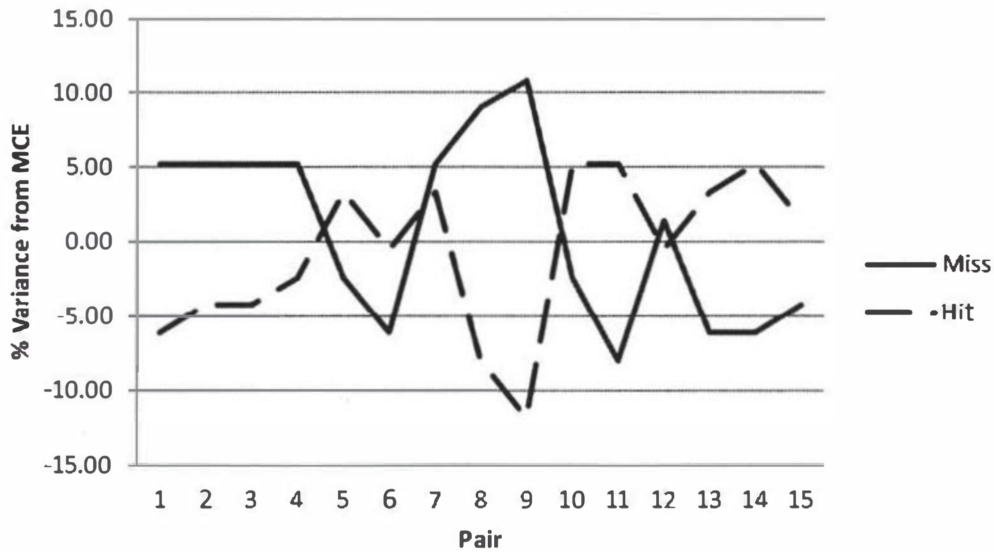

Figure 27. Pair scores for Dataset B.

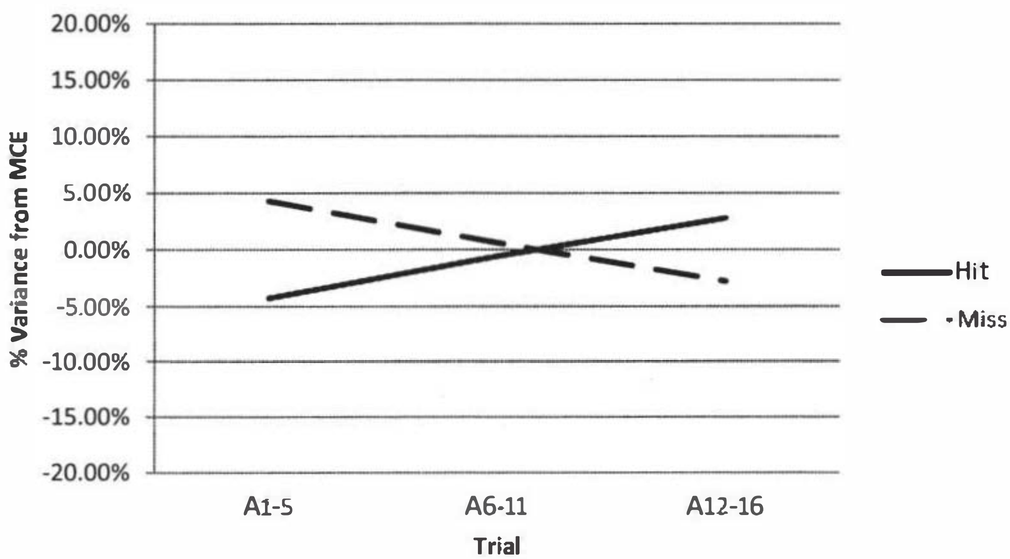

Figure 28. Trial scores split by thirds for Dataset B.

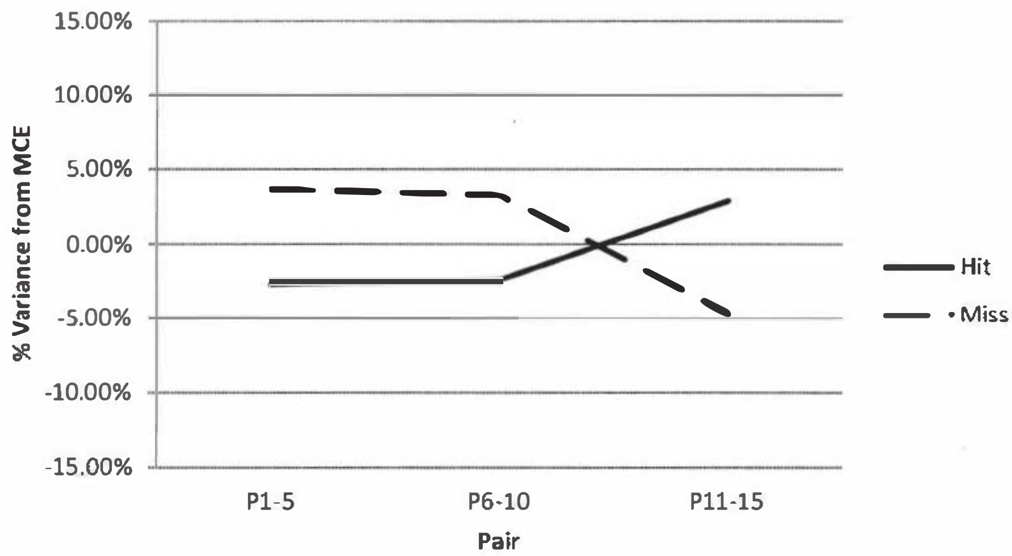

Figure 29. Pair scores split by thirds for Dataset B.

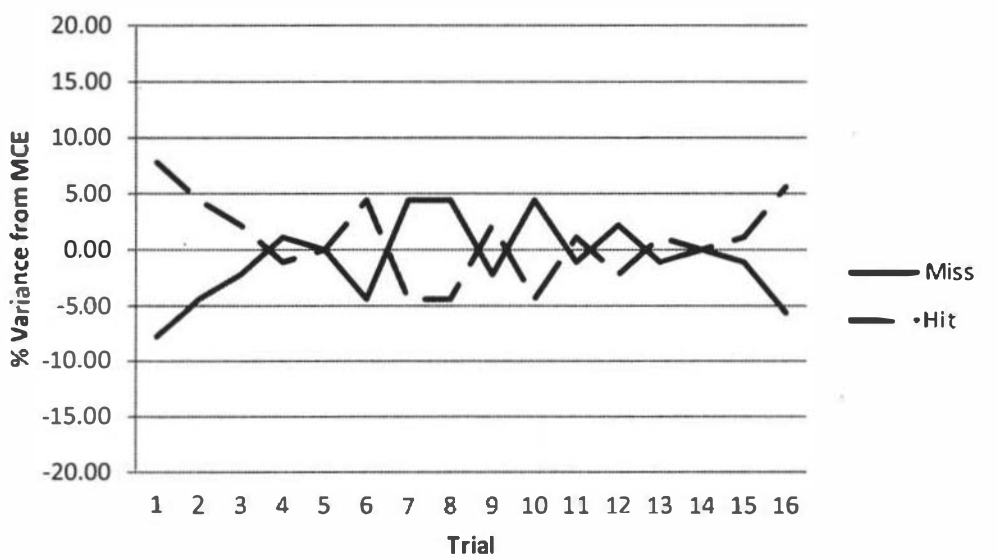

Figure 30. Trial scores for Dataset C.

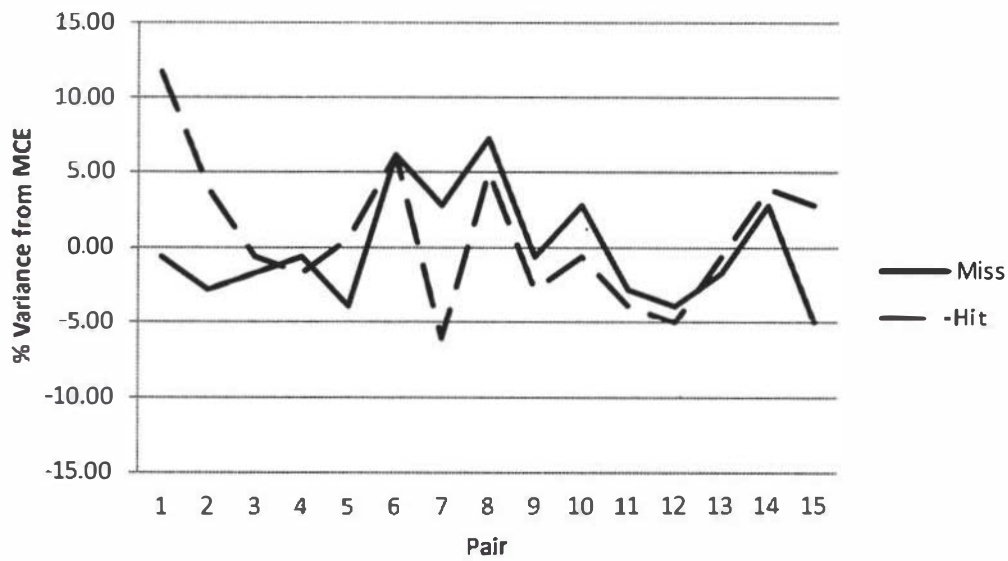

Figure 31. Pair scores for Dataset C.

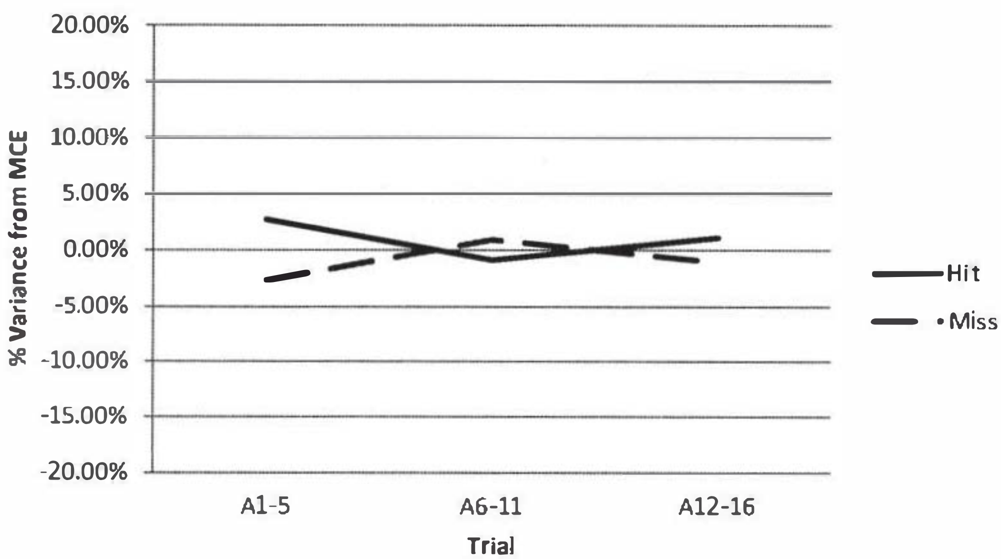

Figure 32. Trials grouped by thirds for Dataset C.

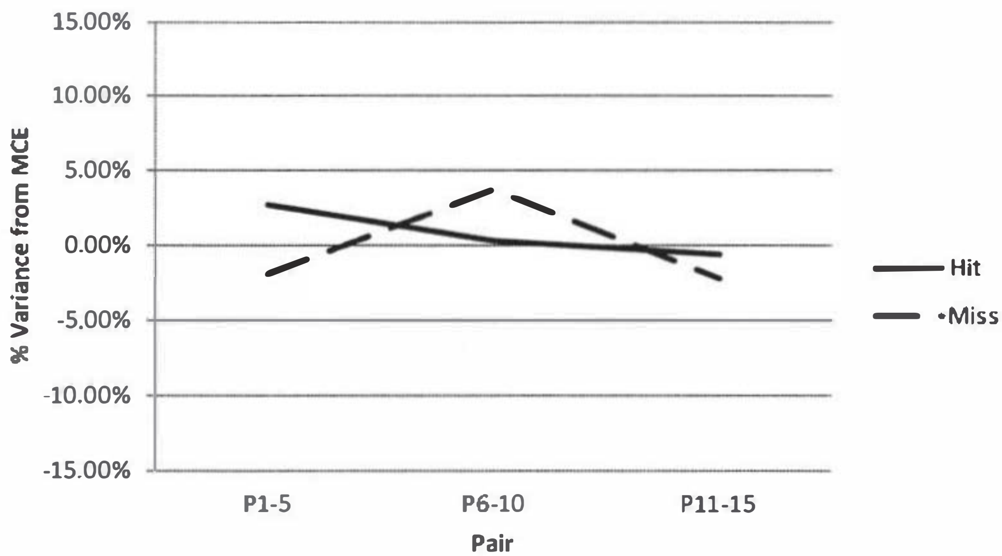

Figure 33. Pairs grouped by thirds for Dataset C.

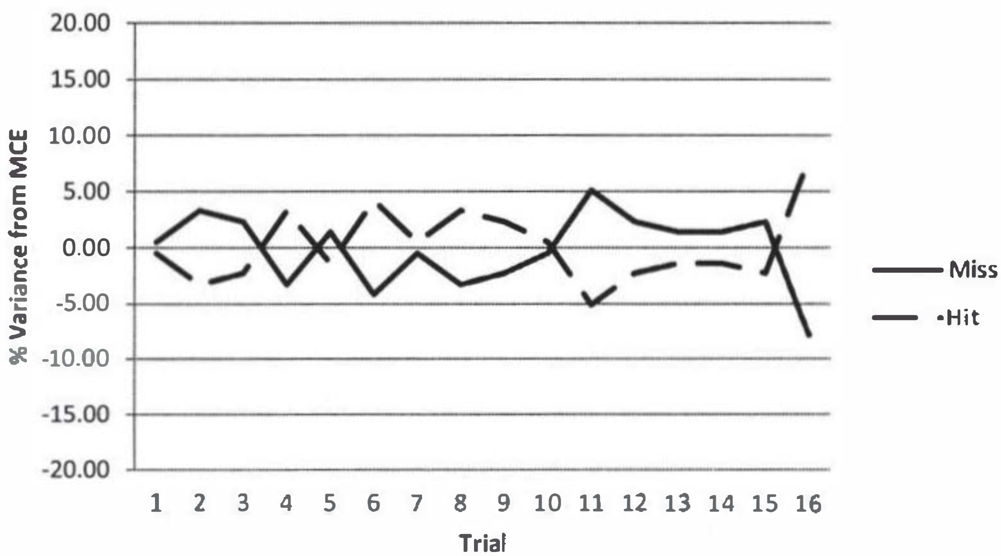

Figure 34. Trial scores for Dataset D.

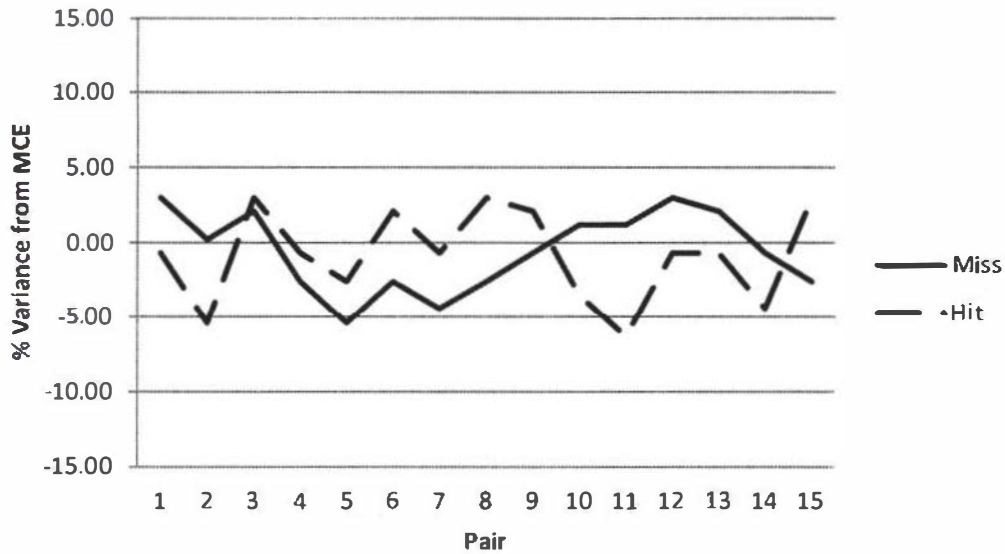

Figure 35. Pair scores for Dataset D.

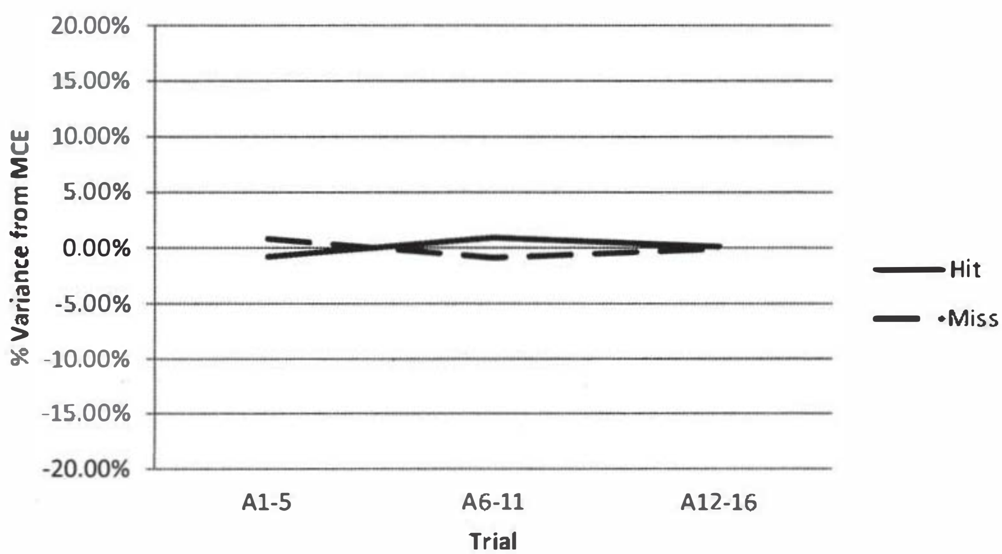

Figure 36. Trial scores grouped by thirds for Dataset D.

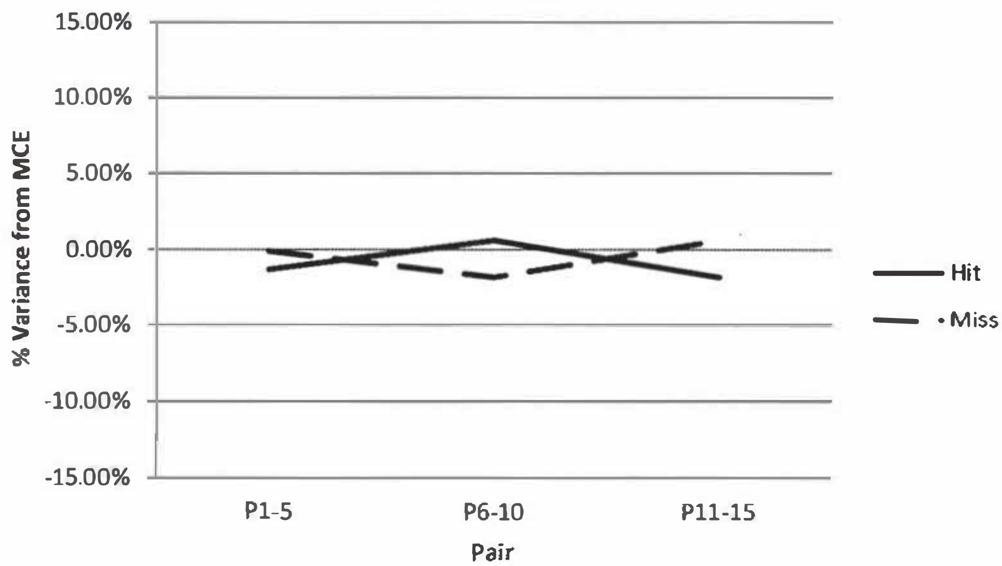

Figure 37. Pair scores grouped by thirds for Dataset D.

References

Anderson, M., & White, R. (1958). A survey of work on ESP and teacher-pupil attitudes. *The Journal of Parapsychology, 22*, 246-268.

Bem, D. J., & Honorton, C. (1994). Does Psi Exist? Replicable Evidence for an Anomalous Process of Information Transfer. *Psychological Bulletin, 115(1),* 4-18.

Blackmore, S. J., & Rose, N. (1997). Reality and imagination: A psi-conducive confusion? *The Journal of Parapsychology, 61,* 321-335.

Bunfill, T. (2006). *Mental Telepathy and Social Fields: A Within Session Decline Effect* (master's thesis). Available from The Keep at Eastern Illinois University (registry number unknown).

Conner, G. M. (2009). *A Three Part Study Exploring Psi-Conduciveness and Psi-Defensiveness in Groups* (master's thesis). Available from The Keep at Eastern Illinois University (registry number unknown).

Crandall, J. E. (1985). Effects of Favorable and Unfavorable Conditions on the Psi-Missing Displacement Effect. *The Journal of the American Society for Psychical Research, 79,* 27-38.

Crandall, J. E., & Hite, D. D. (1983). Psi-Missing and Displacement: Evidence for Improperly Focused Psi? *The Journal of the American Society for Psychical Research, 77,* 209-227.

Fischer, I., & Savranevski, L. (2015). Extending the two faces of subjective randomness: From the gambler's and hot hand fallacies toward a hierarchy of binary sequence perception. *Mem Cogn, 43*, 1056-1070. doi: 10.3758/s13421-015-0523-5

Greenhouse, J. B. (1991). [Replication and Meta-Analysis in Parapsychology]: Comment: Parapsychology--On the Margins of Science?. *Statistical Science, 6(4)*, 386-389. Retrieved from http://www.jstor.org/stable/2245732

Haraldsson, E. (1981). Some Determinants of Belief in Psychical Phenomena. *The Journal of The American Society for Psychical Research, 75(4)*, 297-309.

Hinman, A. J. (2015). *Weird Psi-ence!* (master's thesis). Available from The Keep at Eastern Illinois University (registry number unknown).

Houtkooper, J. M., & Haraldsson, E. (1997). Reliabilities and psychological correlates of guessing and scoring behavior in a forced-choice ESP task. *The Journal of Parapsychology, 61,* 119-134.

Kelley, M. P. (n.d.). Correlates of paranormal beliefs, I: Schizotypy. *The Journal of Parapsychology, unknown volume,* pp. 301-317.

Kelly, E. F. (1980). Further notes on consistent missing. *The Journal of Parapsychology, 44,* 57-61.

Kennedy, J. E. (1979). Consistent missing: A type of information-processing error in ESP. *Journal of Parapsychology, 43*, 113-127.

Kennedy, J. E. (1994). Exploring the limits of science and beyond: Research strategy and status. *Journal of Parapsychology, 58,* 59-77.

Kennedy, J. E. (1995). Methods for investigating goal-oriented psi. *The Journal of Parapsychology, 59,* 47-62.

Luke, D. P., Delanoy, D., & Sherwood, S. J. (2008). Psi may look like luck: Perceived luckiness and beliefs about luck in relation to precognition. *Journal of the Society for Psychical Research, 72(4),* 193-207.

McWhorter, B. (2005). *Social fields, belief in psi, & their effect on mental telepathy* (master's thesis). Available from The Keep at Eastern Illinois University (registry number unknown).

Morrison, R. S., & Ordeshook, P. C. (1975). Rational choice, light guessing and the gambler's fallacy. *Public Choice, 22,* 79-89.

Murphy, K., & Lester, D. (1976). A search for correlates of belief in ESP. *Psychological Reports, 38,* 82.

Nash, C. B. (1976). Group selection of target painting. *European Journal of Parapsychology, 1(3),* 37-39.

Nickerson, R. S., & Butler, S. F. (2009). On producing random binary sequences. *The American Journal of Psychology, 122(2),* 141-151.

Palmer, J. (2000). Covert psi in computer solitaire. *The Journal of Parapsychology, 64,* 195-211.

Parker, A. (2003). We Ask, Does Psi Exist? But is this the right question and do we really want an answer anyway? *Journal of Consciousness Studies, 10(6-7),* 111-134(24).

Parker, J. (2006) *Psi Hitting and Missing: The Effects of Belief, Order of Run, and Previous Experience* (master's thesis). Available from The Keep at Eastern Illinois University (registry number unknown).

Rao, K. R., Kanthamani, H., & Krishna, S. R. (1979). Psi and volitional activity: An experimental study of psi in a competitive situation. *The Journal of Parapsychology, 43,* 101-111.

Rao, K. R., & Kanthamani, H. (1980). ESP tests with competition set: Further explorations of volitional influence on psi. *The Journal of Parapsychology, 44,* 341-351.

Schmeidler, G. R., & Murphy, G. (1945). The influence of belief and disbelief in ESP upon individual scoring levels. Journal unknown. pp. 271-276.

Sheldrake, R. (2014). Telepathy in Connection with Telephone Calls, Text Messages and Emails. *Journal of International Society of Life Information Science, 32(1),* 7-15.

Smith, M. D. (2003). The psychology of the "psi-conducive" experimenter: Personality, attitudes towards psi, and personal psi experience. *The Journal of Parapsychology, 67,* 117-128.

Standish, L. J., Johnson, L. C., Kozak, L., & Richards, T. (2003). Evidence of correlated functional magnetic resonance imaging signals between distant human brains. *Alternative Therapies, 9(1),* 122-128.

Storm, L., Tressoldi, P. E., & Di Risio, L. (n.d.). Meta-analysis of ESP studies, 1987-2010: Assessing the success of the forced-choice design in parapsychology. *The Journal of Parapsychology, unknown volume,* pp. 243-261.

Sundali, J., & Croson, R. (2006). Biases in casino betting: The hot hand and the gambler's fallacy. *Judgment and Decision Making, 1,* 1-12.

Tobacyk, J., & Milford, G. (1983). Belief in Paranormal Phenomena: Assessment Instrument Development and Implications for Personality Functioning. *Journal of Personality and Social Psychology, 44(5),* 1029-1037.

Warner-Angel, H. (2016). *Psi Phenomena: Mental Telepathy in Groups* (master's thesis). Available from The Keep at Eastern Illinois University (registry number unknown).

Watt, C., & Nagtegaal, M. (2000). Luck in action? Belief in good luck, psi-mediated instrumental response, and games of chance. *The Journal of Parapsychology, 64,* 33-52.

Watt, C., & Brady, C. (2002). Experimenter effects and the remote facilitation of attention focusing: Two studies and the discovery of an artifact. *The Journal of Parapsychology, 66,* 49-71.

Watt, C., & Wiseman, R. (2002). Experimenter differences in cognitive correlates of paranormal belief and in psi. *The Journal of Parapsychology, 66,* 371-385.

Watt, C., & Ramakers, P. (2003). Experimenter effects with a remote facilitation of attention focusing task: A study with multiple believer and disbeliever experimenters. *The Journal of Parapsychology, 67,* 99-116.

West, D. (2012). Awkward Questions. *Journal of the Society for Psychical Research, 76(4),* 204-209.

Wiseman, R., & Schlitz, M. (1997). Experimenter effects and the remote detection of staring. *The Journal of Parapsychology, 61,* 197-207.

Xu, J., & Harvey, N. (2014). Carry on winning: The gamblers' fallacy creates hot hand effects in online gambling. *Cognition, 131,* 173-180.

Yarnall, J. A. (2003). *The Effects of Social Fields on the Telepathic Reception of Information* (master's thesis). Available from The Keep at Eastern Illinois University (registry number unknown).