Abstract

This is the first study to examine the relationship between risk intelligence and belief in the paranormal. The risk intelligence test, which was developed by Evans (2012), measures the ability to estimate probabilities, and is based on the established methodology of calibration testing. Approximately 40,000 participants completed the test, and indicated whether they believe in the paranormal or not. As expected, the risk intelligence of skeptics was significantly higher than that of those who believe in the paranormal, even when gender was controlled for. This lends support to the hypothesis proposed by Blackmore (1992; 1997), according to which belief in the paranormal may be favored by a tendency to misjudge probabilities.

Keywords

Risk intelligence; probability; paranormal; calibration; misjudgment; skepticism;

1. Introduction

1.1. Risk Intelligence

The term "risk intelligence" has been defined in various different ways (Apgar 2006, Funston and Wagner 2010), but in this paper we define it as "the ability to estimate probabilities accurately" (Evans, 2012).

How do we judge the accuracy of probability estimates? One way is to compare subjective probability estimates to objective statistics. For example, one can ask people to estimate the probability of death from various causes for some particular demographic group, and compare these estimates to the mortality data. This method is restricted, of course, to subject areas for which data are readily available.

Another way to measure a person's ability to provide accurate probability estimates is calibration testing (Lichtenstein, Fischhoff et al. 1982). This involves collecting many probability estimates whose correct answer is known or will shortly be known to the experimenter, and plotting the proportion of correct answers against the subjective estimates. For example, suppose that every day you estimate the probability that it will rain in your neighborhood the following day, and then you note whether or not it did, in fact, rain on each day. To simplify things a little, let's assume that you can only choose from a discrete set of probability values, such as 0, 0.1, 0.2, etc. Over the course of a year, you collect 365 estimates, for each of which you have also indicated whether it did, in fact, rain or not. Suppose that you estimated the chance of rain as 0 on 15 days. If you are well calibrated, it should have rained on none of those days. Again, if there were 20 days which you assigned a 0.1 probability of rainfall, it will have rained on 2 of those days if you are well calibrated. When the proportions of correct answers are plotted against the subjective estimates, the result is known as a calibration curve. With perfect risk intelligence, all points in the calibration curve would fall on the identity line (x = y). Nobody is perfectly calibrated, but as may be seen from Figure 1, US weather forecasters are pretty close. However, as the same figure also shows, not all professionals are so well calibrated.

It may be objected that these two approaches to judging the accuracy of probability estimates do not measure the same thing. Comparing subjective probability estimates to statistics may appear, *prima facie*, to target a different psychological process than estimating the chances of a single event, such as the chance that it will rain next Tuesday. We dispute this view, but we recognize that it may derive from a different understanding of probability to the one that informs this paper. The concept of probability is the focus of fierce debate between “subjectivists,” who take probabilities to express degrees of belief, and “frequentists,” who take probabilities to refer to objective facts about the world. For the subjectivist, the statement that “there is a 50 percent chance of this coin landing on heads” is an expression of his or her own uncertainty. For the objectivist, however, the statement does not have anything to do with anyone’s beliefs; rather, it means that, in the long run, the coin will land heads up half of all the times it is tossed. We start from a subjectivist approach to probability, according to which there is no such thing as a “true” probability, in the sense of some objective fact existing out there in the world; probabilities are just ways of quantifying our subjective degree of belief. So while we often talk loosely of “making accurate probability estimates,” strictly speaking, this phrase is incoherent. The accuracy of an estimate can be measured only by comparing it to some objective fact, and such facts do not exist in the case of probabilities. This is why experts who study risk intelligence usually prefer to speak of “well-calibrated” probability estimates rather than of accurate ones. A more precise definition of risk intelligence would therefore be: the ability to provide well calibrated probability estimates.

Figure 1 displays data from two different studies. When the weather forecasters in the first study (Murphy & Winkler 1977) said there was a 90 percent chance of rain the following day, it rained almost 90 percent of the time. But when the doctors in the second study (Christensen-Szalanski & Bushyhead 1981) estimated that there was a 90 percent chance that their patients had pneumonia, only about 15 percent of the patients turned out to have the disease. In other words, the doctors in this study had much more faith in the accuracy of their diagnoses than was justified by the evidence. That meant they were likely to recommend more tests than was strictly necessary, prescribe more treatments than warranted, and cause their patients needless worry.

Between 1960 and 1980, psychologists measured the risk intelligence of many specific groups, such as medics and weather forecasters, but did not gather extensive data on the risk intelligence of the general public. One reason for this was no doubt because the testing was done with pen and paper, which made data collection and processing a time-consuming process. It appears that interest in calibration testing began to decline after 1980, and has not progressed much since then. This area of research is ripe for revival, especially now that the internet allows testing and data collection to be automated.

In this paper, we present data on the risk intelligence of a sample of over 10,000 people of all ages and from a wide variety of countries. We were able to collect such a large amount of data by using an online calibration test rather than a pencil and paper version.

1.2. Paranormal Beliefs

Paranormal beliefs encompass a range of beliefs in phenomena that "transcend the explanatory power of mainstream science" (Gray 1991:7). They may be classified into two groups: religious paranormal beliefs, which include beliefs central to the religious doctrines (Rice 2003), and classical paranormal beliefs, which comprise occult and supernatural beliefs (Irwin 2009). The first group traditionally include: belief in God, belief in life after death, belief in heaven and hell, and belief in the devil (Aarnio & Lindeman, 2007; Hillstrom 2000; Rice 2003). On the other hand, paranormal beliefs include the traditional paranormal powers of extra-sensory perception and psycho-kinesis (Wiseman & Watt 2006) as well as such diverse phenomena as déjà vu, ghosts, witchcraft, good luck omens, psychic healing, reincarnation, communication with the dead, astrology and extraordinary life forms (Aarnio & Lindeman, 2007; Hillstrom 2000; Irwin 2009; Rice 2003; Tobacyk 2004).

1.3. Probability Misjudgment theories

Why do some people believe in the paranormal, while others are more sceptical? Blackmore (1992; 1997; Blackmore & Troscianko; 1985) suggested that paranormal believers are less competent than non-believers in making accurate probability judgments. She argued that if people underestimate the probability of chance coincidences they will be surprised when they occur, will seek an explanation and so be more likely to misinterpret chance events as evidence of the paranormal. To test this hypothesis, Blackmore and Troscianko (1985) asked participants to perform several computer controlled probability tasks, including generating a random string of numbers, assessing whether samples were biased, coin tossing, and sampling with and without replacement. Participants were then assessed for their belief in the paranormal. As expected, believers underestimated the number of doubles in random sequences (the same digit produced twice consecutively) more than non-believers. Indeed the believers performed worse than non-believers in all the tasks.

Subsequent research has extended these findings, testing the avoidance of repetitions when generating strings of random numbers (Bressan, 2002; Brugger, 1997; Brugger, Landis, & Regard, 1990; Brugger, Regard, & Landis, 1991; Brugger, Regard, Landis, Krebs, & Niederberger, 1994) using variations on the ‘birthday paradox’ (Matthews &Blackmore, 1995) and measuring susceptibility to the ‘conjunction fallacy’; that is to misperceiving conjunct events as being more likely than singular events alone (Rogers, Davis & Fisk 2009). Although some studies have failed to support the hypothesis (e.g. Houtkooper & Haraldsson, 1997; Matthews & Blackmore, 1995), the findings generally support it.

One possibility is that these effects are due to differences in general cognitive ability, as argued by Musch and Ehrenberg (2002). They gave a battery of probabilistic reasoning tasks to 123 students and found a significant correlation between error rates and paranormal belief, but this correlation disappeared when cognitive ability, measured by final exam grades, was controlled for. However, their sample was small and there are other reasons to doubt this conclusion. For example, Pennycook, Cheyne, Seli, Koehler, and Fugelsang, (2012) showed that a correlation between paranormal belief and analytic cognitive style (a propensity to question and set aside highly salient intuitions) did not disappear when controlled for general cognitive ability. In addition, the many studies that have looked for correlations between cognitive ability and paranormal belief have produced highly inconsistent findings. In a review of the literature, Wiseman and Watt (2006) conclude that “unlike the literature on general cognitive functioning, the existing literature provides some evidence to support the notion that disbelievers tend to outperform believers on tasks involving probabilistic judgments." (Wiseman & Watt 2006, p: 329).

The purpose of the present study is to extend this research and to look at risk intelligence, which does not appear to depend on reasoning skills but on intuitions about probabilities (Ceci & Liker, 1986). If Blackmore's hypothesis is correct, we would expect to find a negative correlation between risk intelligence and belief in the paranormal.

2. Material and methods

2.1. Participants

Between 1 January and 31 December 2010, we collected data from 37,943 respondents, of whom 31,808 (83.8%) answered the question about the paranormal. We excluded those respondents who did not answer all 50 questions in our test or did not specify their sex or did not give consent for their data to be used in our research. After these exclusions, 10,802 respondents remained in our sample.

Almost 80% of the participants were male (8,520). The participants were citizens from over 20 countries distributed over the five continents. The most well-represented countries were the USA, the UK, Ireland, Canada, Australia, and Germany. Participants ranged in age from under 10 to over 80, though most were aged between 21 and 40.

2.2. Procedure

The participants were informed about the study through blogs, press releases, media interviews and Internet discussion forums. The response rate cannot be reliably calculated since no data are available on how many people were reached by these recruitment means.

The test was available on the Internet through a website (www.projectionpoint.com). The website was hosted on a Linux server operated by Blacknight Hosting. The test itself is written in PHP using JavaScript as the client-side scripting language, and all data were stored in a MySQL database. If participants provided more than one estimate for any statement, only the first estimate provided was stored. Participants were not asked to provide private data such as names, addresses or telephone numbers, and they indicated their age by specifying a range, so their exact ages remain unknown. Later versions of the test (representing around 90% of the subjects) stored IP addresses and browser data; however, none of these data were used when analyzing the test results, since the information provided by such data was not relevant to our study. Upon beginning a test, subjects were assigned a unique integer ID - this ID allowed us to uniquely but anonymously distinguish between participants.

The test was designed to be supported by all major browsers (Internet Explorer, Firefox, Chrome, Opera and Safari). The approximate response time of the test was around 9 minutes.

2.3. Measures

2.3.1. Risk Intelligence Measure. Risk Intelligence (RQ) was measured through the online test developed by Evans (2012). It consisted of 50 statements about general knowledge, such as “A one followed by 100 zeros is a Googol,” and “Africa is the largest continent” (see Appendix for full list of statements). The participants were asked to indicate how likely they thought it was that each statement was true by clicking on one of eleven buttons that indicated percentage values ranging from 0% to 100% in increments of 10%. If the participant was absolutely sure that a statement is true, they were instructed to click on the 100% button. Conversely, they were instructed to click on the 0% button if they were convinced that a statement was completely false. In cases where they thought the statement was probably true, but were not completely sure, they were asked to click on the 60%, 70%, 80%, or 90% buttons, depending on how sure they were, and likewise for the 40%, 30%, 20%, and 10% buttons if they thought the statement was probably false. They were asked to click on the button marked 50% if they were completely uncertain as to whether a statement was true or false. When a participant had estimated the likelihood of all fifty statements in the test, the website calculated their risk intelligence quotient, or RQ, a number between 0 and 100, according to the algorithm detailed in Appendix B.

2.3.2. Belief in the Paranormal Measure. There are numerous ways of measuring paranormal belief (Irwin, 2009) but most ask about a range of phenomena irrelevant to the hypothesis being considered here. The most widely used is probably Tobacyk's (2004) revised paranormal scale, which probes seven factors and includes witches, God, and communication with the dead. The Australian Sheep-Goat Scale (Thalbourne, 1995; Thalbourne, Dunbar &Delin 1995) is also widely used and includes life after death. Also some items of these scales can be misleading to some participants. For example, people may be unsure whether the question "Do you believe in astrology?" refers to their conviction in the ability of horoscopes to predict the future or to reveal their personality. The use of a better-designed paranormal belief scale may provide a promising avenue for further research, but in the current study, the "belief in the paranormal" variable was simply assessed through the direct binary question: "Do you believe in the paranormal?"

3. Results

Tables 1 and 2 show the distribution of belief in the paranormal in our sample by sex. Men outnumbered women in our sample by a factor of four, and skeptics outnumbered believers in the paranormal by more than a factor of five. The ratio of skeptics to believers in the men was more than double that in the women.

Because of these sex differences, we analyzed the risk intelligence scores separately for men and women. The results of this analysis are shown in Table 3. It may be seen that in both men and women, the risk intelligence of skeptics was significantly higher than that of those who believe in the paranormal.

4. Discussion

To the best of our knowledge, this is the first study that examines the correlation between risk intelligence and the belief in the paranormal. In addition to being the first of its kind, this study also involved a sample size that is several orders of magnitude larger than all previous studies involving calibration testing.

People with high risk intelligence tend to judge probabilities accurately (Evans, 2012). According to Blackmore, those who believe in the paranormal are likely to misjudge probabilities (Blackmore 1985; 1997). Accordingly, we hypothesized a negative correlation between risk intelligence and belief in the paranormal. Our findings are consistent with this hypothesis, since skeptics scored significantly higher than paranormal believers on the risk intelligence test developed by Evans (2012). This is true for both men and women when our data are analyzed separately by sex.

Our results are consistent with a broader framework for understanding the differences between those who are skeptical about and those who believe in the paranormal which has emerged over the past decade. This framework attributes some of the variation in paranormal belief to variations in the tendency to perceive patterns in ambiguous or statistically random data (Blackmore & Moore, 1994). This variation has in turn been linked to a specific dopaminergic genetic polymorphism (COMT) known for its involvement in prefrontal executive cognition (Raz, Hines, Fossella & Castro, 2008). According to this model, people with high levels of dopamine are more likely to find significance in coincidences, and pick out meaning and patterns where there are none, which would give rise to both lower risk intelligence and a greater tendency to believe in the paranormal (Krummenacher, Mohr, Haker, & Brugger, 2010). To the best of our knowledge, no studies have investigated this causal hypothesis directly. It would, therefore, be an interesting direction for future research.

Another possible direction for future research would be to use Tobacyk's revised paranormal belief scale to distinguish between various aspects of paranormal belief and discover which of these account for the correlation with risk intelligence (Tobacyk, 2004). The seven factors are Traditional Religious Belief, Psi, Witchcraft, Superstition, Spiritualism, Extraordinary Life Forms, and Precognition. We predict that Psi, Superstition and Precognition would jointly account for all the correlation because these apply to the original hypothesis about probabilities. The other factors do not involve an element of probability.

In the current study, the "belief in the paranormal" variable was simply assessed through the direct binary question: "Do you believe in the paranormal?" The use of a better-designed paranormal belief scale would provide a promising avenue for further research, permitting the detection of a more fine-grained correlation between degrees of risk intelligence and degrees of belief in the paranormal. Even formulating our simple question in a Likert-type way to obtain a gradient of belief may provide some useful data in this regard.

Finally, it would also be interesting to enquire whether training programs designed to improve risk intelligence also tend to reduce belief in the paranormal. It has been suggested that risk intelligence may be improved given repeated practise in estimating probabilities, providing that trainees are provided with prompt and well-defined feedback (Lichtenstein, Fischhoff & Phillips, 1982). Given the mixed results with such training programs so far, and the disappointing results of previous interventions aimed at reducing paranormal belief, any positive findings in this area would be highly significant.

References

Aarnio, K., & Lindeman, M (2007). "Religious people and paranormal believers alike or different." *Journal of Individual Differences 28*(1): 1-9.

Apgar, D. (2006). Risk Intelligence: Learning to Manage What We Don't Know. Cambridge, Mass., Harvard Business School Press.

Blackmore (1997). "Probability misjudgment and belief in the paranormal : A newspaper survey." *The British Psychological Society, 88*, 683-689.

Blackmore, SJ 1992 Psychic Experiences: Psychic Illusions *Skeptical Inquirer* 16 367-376.

Blackmore,S.J. and Moore,R. (1994). Seeing things: Visual recognition and belief in the paranormal. *European Journal of Parapsychology*, 10, 91-103.

Blackmore, S.J. and Troscianko, T.S. (1985). "Belief in the Paranormal: Probability judgements, illusory control and the "chance baseline shift". *British Journal of Psychology, 76,* 459-468.

Bressan, P. (2002). The connection between random sequences, everyday coincidences, and belief in the paranormal. *Applied Cognitive Psychology*, 16, 17–34.

Brugger, P. (1997). Variables that influence the generation of random sequences. *Perceptual and Motor Skills*, 84, 627–661.

Brugger, P., Landis, T. & Regard, M. (1990). A sheep-goat effect in repetition avoidance: Extrasensory perception as an effect of subjective probability. *British Journal of Psychology, 81,* 455-468.

Brugger, P., Regard, M., & Landis, T. (1991). Belief in extrasensory perception and illusory control: A replication. *Journal of Psychology*, *125*, 501–502.

Brugger, P., Regard, M., Landis, T., Krebs, D., & Niederberger, J. (1994). Coincidences: Who can say how meaningful they are? In E. W. Cook & D. L. Delanoy (Eds.), *Research in parapsychology 1991,* Metuchen, NJ: Scarecrow Press. 94–98

Ceci, S. & Liker, J. (1986) A Day at the Races: A Study of IQ, Expertise, and Cognitive Complexity, *The Journal of Experimental Psychology: General 115*, 255–266.

Christensen-Szalanski, J. J. J. and J. B. Bushyhead (1981). "Physicians' use of probabilistic information in a real clinical setting." *Journal of Experimental Psychology: Human Perception and Performance, 7*, 928-935.

Evans, Dylan (2012). Risk Intelligence: How to Live with Uncertainty. New York, Free Press.

Funston, F. and S. Wagner (2010). Surviving and Thriving in Uncertainty: Creating the Risk Intelligent Enterprise. Hoboken, NJ, Wiley.

Gray, W. D. (1991). Thinking critically about New Age ideas. Belmont, CA: Wadsworth.

Hillstrom, E. L., & Strachan, M. (2000). "Strong commitment to traditional protestant beliefs is negatively related to beliefs in paranormal phenomena." *Psychological Reports, 86,* 183-189.

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

Irwin, H.J. (2009). The Psychology of Paranormal Belief: A Researcher's Handbook University of Hertfordshire Press.

Krummenacher, P., Mohr, C., Haker, H., & Brugger, P. (2010). Dopamine, paranormal belief, and the detection of meaningful stimuli. *Journal of Cognitive Neuroscience*, *22*(8), 1670-1681. doi: 10.1162/jocn.2009.21313

Lichtenstein, S., B. Fischhoff, Phillips, L.,D. (1982). Calibration of probabilities: The state of the art to 1980. Judgement Under Uncertainty: Heuristics and Biases. D. Kahneman, P. Slovic and A. Tversky. Cambridge, Cambridge University Press: 306-334.

Matthews, R., & Blackmore, S. J. (1995). Why are coincidences so impressive? *Perceptual and Motor Skills*, *80,* 1121–1122.

Murphy, A. H. and R. L. Winkler (1977). "Can weather forecasters formulate reliable probability forecasts of precipitation and temperature?" National Weather Digest 2(May): 2-9.

Murphy, A. H. and R. L. Winkler (1977). "Reliability of subjective probability forecasts of precipitation and temperature." Journal of the Royal Statistical Society. Series C (Applied Statistics) 26(1): 41-47.

Musch,J. and Ehrenberg,K. (2002). Probability misjudgement, cognitive ability, and belief in the paranormal, *British Journal of Psychology, 93*, 169-177.

Pennycook, G., Cheyne, J.A., Seli, P., Koehler, D.J. and Fugelsang, J.A. (2012). Analytic cognitive style predicts religious and paranormal belief, *Cognition*, *123*, 335-346.

Raz, A., Hines, T., Fossella, J., & Castro, D. (2008). Paranormal experience and the comt dopaminergic gene: A preliminary attempt to associate phenotype with genotype using an underlying brain theory.*Cortex*, *44*, 1 3 3 6 – 1 3 4 1. doi: 10.1016/j.cortex.2007.07.011

Rice, T. W. (2003). "Believe it or not: Religious and other paranormal beliefs in the united states." *Journal for the Scientific Study of Religion* , *42(1)*, 95-106.

Rogers, P., Davis, T. and Fisk, J. (2009). Paranormal belief and susceptibility to the conjunction fallacy. *Applied Cognitive Psychology*, *23*, 524-542

Thalbourne, Michael A. (1995). Further studies of the measurement and correlates of belief in the paranormal. *Journal of the American Society for Psychical Research*, *89(3),* 233-247.

Thalbourne, M., Dunbar, K.A. and Delin, P.S. (1995). An investigation into correlates of belief in the paranormal, *Journal of the American Society for Psychical Research*, *89(3),* 215-231.

Tobacyk, J.J. (2004). A revised paranormal belief scale, *International Journal of Transpersonal Studies, 23,* 94-98

Wiseman, R. , Watt,C. (2006). Belief in psychic ability and the misattribution hypothesis: A qualitative review, *British Journal of Psychology*, *97*, 323–338

Appendix A - Risk Intelligence Test (Evans, 2012)

The fifty true/false statements in the calibration test were as follows:

A one followed by 100 zeros is a Googol	T
Africa is the largest continent	F
Alzheimer's accounts for under half the cases of dementia in the US	F
An improper fraction is always less than one	F
Armenia shares a common border with Russia	F
There have been over 40 US Presidents	T
In 1994, Bill Clinton was accused of sexual harassment by a woman called Paula Jones	T
Canberra is the capital of Australia	T
Cats are not mentioned in the Bible	T
Christianity became the official religion of the Roman empire in the third century AD	F
Commodore Matthew Perry compelled the opening of Japan to the West with the Convention of Kanagawa in 1870	F
El Salvador does not have a coastline on the Caribbean	T
Gout is known as "the royal disease"	F
Harry Potter and the Goblet of Fire tells the story of Harry Potter's third year at Hogwarts	F
Humphrey Bogart had two wives before Lauren Bacall	F
In 2008 the population of Beijing was over 20 million people	F
In the Old Testament, Jezebel's husband was Ahab, King of Israel.	T
Iron accounts for over 30 percent of the Earth's composition	T
It is possible to lead a cow upstairs but not downstairs, because a cows' knees can't bend properly to walk back down	T
Lehman Brothers went bankrupt in September 2008	T
LL Cool J got his name from the observation "Ladies Love Cool James"	T
Male gymnasts refer to the pommel horse as "the pig"	T
Mao Zedong declared the founding of the People's Republic of China in 1949	T
More than 10 American states let citizens smoke marijuana for medical reasons	T
More than 8 out 10 victims infected by the Ebola virus will die in two days	T
Most of the terrorists who carried out the attacks on 9/11 were from Saudi Arabia	T
Mozart composed over 1000 works	F
Natural gas has an odor	F
Of all Arab nations, Lebanon has the highest percentage of Christians	T
Over 40 per cent of all deaths from natural disasters from 1945 to 1986 were caused by earthquakes	T
Over 50 % of Nigeria's population lives on less than one dollar per day	T

Stalagmites grow down, and stalactites grow up	F
The Italian musical term adagio means that the music should be played quickly	F
The Euphrates river runs through Baghdad	F
The face on a $100,000 bill is that of Woodrow Wilson.	T
The Islamic Resistance Movement is better known to Palestinians as Hizbollah	F
The Japanese were largely responsible for building most of the early railways in the U.S. West	F
The last Inca emperor was Montezuma	F
The most frequently diagnosed cancer in men is prostate cancer	T
The only stringed symphonic instrument that has a pedestal and a crown is a double bass	F
The president of Russia is Vladimir Putin	F
The San Andreas Fault forms the tectonic boundary between the Pacific Plate and the North American Plate.	T
The US civil war broke out the same year the federal government first printed paper money	T
The US Declaration of Independence begins: "We the People of the United States..."	F
The word 'robot' was coined by the American science fiction writer, Isaac Asimov	F
The world's highest island mountain is Mauna Kea	T
The Taj Mahal was built by Emperor Shah Jahan in memory of his favorite wife	T
There are more people in the world than chickens	F
There are no diamond fields in South America	F
Wikipedia was launched in 1999 by Jimmy Wales and Larry Sanger	F
NUMBER OF TRUE STATEMENTS	25

Appendix B

We used the following algorithm to score the risk intelligence test:

1. Start by counting all the times you assigned a likelihood of 0 percent to a statement, and then count how many of those statements were actually true (the truth values of each statement from the test are below). Then divide the former into the latter and express the answer as a percentage. For example, if there are five statements that you estimated had a 0 percent chance of being true and exactly one of these statements was true, divide five into one, which is 0.2 (or 20 percent). Since you can't divide by zero, if none of the statements was true, just put 0 percent.
2. Do the same for each of the other categories (10 to 100 percent).
3. Find the difference between each of the results you have calculated so far and the value of that category. For example, if, 20 percent of the statements to which you assigned a probability of 0 percent were actually true, the difference is 20. If 30 percent of the statements to which you assigned a probability of 20 percent were actually true, the difference is 10. These are the "residuals."
4. Subtract each residual from 100.
5. Multiply the results from step 4 by the number of times you used the relevant category. For example, if the residual of the 20 percent category is 10 and you assigned a probability of 20 percent to seven statements, multiply 90 by 7.

6. Add up the results from step 5.
7. Divide the result of step 6 by the total number of probability estimates. If you answered all the questions in the test, it is the same as dividing the result of step 6 by the number of questions. This is the weighted mean.
8. Find the square of the result from step 7, and divide it by 100. This is your RQ score.

We chose to use this approach, rather than the better-known Brier score, for three reasons. Firstly, our approach is easier to understand for a lay audience. Secondly, the Brier score is a composite measure of calibration, resolution, and knowledge, whereas we wish to measure only calibration. Finally, we find some of the statistical properties of the Brier score to be unsatisfactory. For example, the Brier score places a premium on extreme forecasts, so that a 100 percent forecast of rainfall is rewarded more when it does, in fact, rain than is a 90 percent forecast of rainfall. That may make sense in the context of weather forecasting, which is what this scoring method was designed for, since it may be the case that in this context “the most useful forecasts are those which fall into the extreme classes,” as Brier argued. But when estimating probabilities in general we see no particular value in valuing more extreme probability estimates more highly than intermediate ones. Even Brier’s remark about the greater usefulness of extreme forecasts may be valid only for populations with low risk intelligence. When people are more comfortable with intermediate probabilities and know how to incorporate probabilistic information into their decisions, extreme probability estimates should be no more useful than intermediate ones.

Ethical Approval

Ethical approval for this research was granted by the Social Research Ethics Committee of University College Cork.

Disclosure of Interests

Dylan Evans and Benjamin Jakobus are shareholders in Projection Point Limited, which provides risk intelligence tests to corporate clients. In addition, Evans is also a Director of the company. This company was incorporated in Ireland in February 2011, after the data reported in this paper had been collected. Throughout the duration of this study, the website www.projectionpoint.com was a purely non-profit research portal, funded entirely by Evans. Only after the study had concluded did Evans and Jakobus decide to pursue commercial applications of risk intelligence, and redesign the website along commercial lines.

Figure 1: Calibration curves from two studies. The first study (solid line) shows data for the US weather forecasters in the study by Murphy and Winkler (1977) . The second study (dashes) shows data for the doctors in the study by Christensen-Szalanski and Bushyhead (1981).

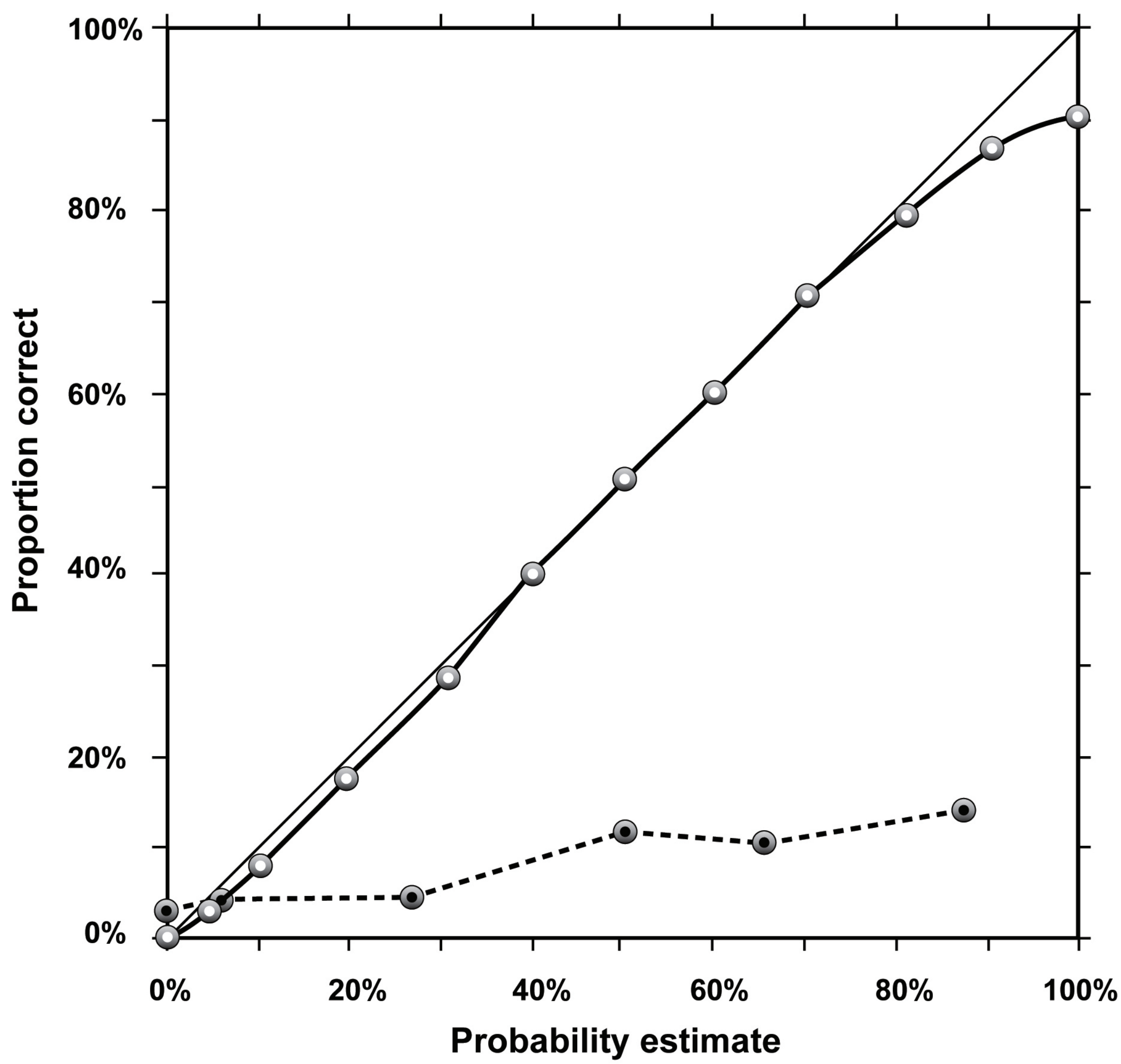

Table 1

Belief in the paranormal by sex

	Believe in paranormal	Do not believe in paranormal	Total
Men	1102 (13%)	7418 (87%)	8520
Women	561 (25%)	1721 (75%)	2282
Total	1663 (15%)	9139 (85%)	10802

Table 2

Chi-square tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	187.533[a]	1	.000		
Continuity Correction[b]	186.640	1	.000		
Likelihood Ratio	170.909	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	187.516	1	.000		
N of Valid Cases	10802				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 351.32.

b. Computed only for a 2x2 table

Table 3

Risk intelligence differences between believers and nonbelievers in the paranormal

	Paranormal Yes		Paranormal No		Mean Diff	95% CI of the difference		P-Value*
	mean RQ	stdev	mean RQ	stdev		Lower	Upper	p
Male	57.12	16.53	64.11	16.32	6.99	5.95	8.02	< .00001
Female	55.65	16.82	61.64	16.97	5.99	4.38	7.60	< .00001

*The p-value was obtained using the t- test