

TITLE PAGE

**'Twitter' as a new research tool:
A mass participation test of remote viewing**

Prof Richard Wiseman
University of Hertfordshire
Dr Caroline Watt
University of Edinburgh

Corresponding author:
Dr Caroline Watt
Department of Psychology
School of Philosophy, Psychology and Language Sciences
University of Edinburgh
7 George Square
Edinburgh EH8 9JZ
UK

Email: Caroline.Watt@ed.ac.uk
Phone: 0131 650 3382 (work)
Phone: 077909 43128 (mobile)

Running title: Twitter Remote Viewing Study

‘Twitter’ as a new research tool: Proof of principle with a mass participation test of remote viewing

Abstract

The social networking site ‘Twitter’ was used to conduct a mass participation remote viewing ESP study. The easy accessibility of Twitter made it possible to recruit and engage a large number of participants, and to give them almost immediate feedback. A majority voting technique was used to combine participants’ calls, to avoid stacking effects and to detect any group-level psi effect. For each trial an experimenter visited the target location. Blind judging was conducted with photographs of the target location and four decoy locations. Over five thousand responses were gathered over five trials. The first trial employed a non-blind judging procedure to test the hypothesis that believers would be especially likely to exhibit confirmation bias. As predicted, a significant relationship was found between belief in psychic ability and level of perceived correspondence between the participants’ impressions and the target location. The following four trials used blind judging. On each trial the group failed to identify the correct target. There was no significant relationship between belief in psychic ability and choice of target on any of the trials. Participants reporting a strong belief in psychic ability identified the correct target on one trial (exact binomial $p=.41$). Those participants who reported that they believed they were psychic and were confident of their response failed to identify the correct target on any trial.

Introduction

Since the 1930s, parapsychologists have used mass participation methods in an attempt to increase the statistical power of their studies and address the question of small effect sizes (Goodfellow, 1938; Milton & Wiseman, 1999). However, typically mass participation studies involve participants sitting at home trying to guess the identity of a distant target, and not receiving feedback until days or weeks later, something that is quite unlike laboratory ESP methods, and not at all engaging for participants. Today, the explosion of Internet technologies presents researchers with improved opportunities not only to involve large numbers of participants, but also to engage them with immediate feedback. The work described here

presents the first parapsychology study to recruit and run participants using the popular social networking tool 'Twitter'.

Twitter is an Internet-based micro-blogging system that allows an individual to instantly send text-based messages of up to 140 characters (known as 'tweets') to a group of people that have chosen to 'follow' that individual (see <https://twitter.com/>). Since its creation in 2006, Twitter has gained considerable attention from the online community, attracted millions of users, and is consistently ranked as one of the world's most visited websites. To discover whether Twitter could be used as an effective research tool, this paper reports a proof of principle study exploring a form of alleged psychic ability known as 'remote viewing'. The study also explored reasons for participants' subjective perceptions of correspondence between target locations and their impressions.

Remote viewing is the alleged ability to gather information about a distant location using paranormal means. The best-known work into this apparent phenomenon was initiated by the American government in the early 1970s, and involved over twenty years of experimental and applied research exploring whether it could assist military intelligence (Mumford, Rose & Goslin, 1995). Some of this experimental research was examined by two reviewers on behalf of the American Institutes for Research (AIR) in the mid 1990s, with one reviewer concluding that the work had demonstrated the existence of a statistically significant effect (Utts, 1995) and the other arguing that it had not been independently replicated (Hyman, 1995).

In remote viewing research individuals are asked to try to psychically identify a distant target location, with their success being determined by the level of correspondence between their thoughts and the location. Skeptics have criticised some of the work, noting that any non-blind assessment of results is open to confirmation bias, with people overestimating the level of correspondence between their thoughts and a target by ignoring mismatches and interpreting any ambiguous elements in their favour (Marks, 2000). It has been further hypothesised that individuals who believe in psychic ability may be more especially likely to demonstrate this bias (Blackmore, 1992; Brugger & Graves, 1997), with indirect support for this notion coming from studies showing that believers are significantly more able than others to 'find' non-existent images within random dot patterns (Blackmore & Moore, 1994; Brugger et al., 1993), and illusory connections between randomly paired stimuli (Brugger, Regard, Landis, Krebs, & Niederberger, 1994; Wiseman & Smith, 2002; Pizzagalli, Lehmann, & Brugger, 2001). Only

one previous study has addressed the issue somewhat more directly by presenting believers and skeptics with made-up data from a fictitious remote viewing study (French, Herrmann, Hales, & Northam, 1996). As predicted, the believers reported finding especially high levels of correspondences between the remote viewers' comments and the alleged target.

The first part of the current study was the first to employ a non-blind judging procedure during a genuine test of remote viewing to directly test the hypothesis that believers would be especially likely to exhibit confirmation bias. An experimenter travelled to a distant location and then asked participants to tweet their thoughts about the nature of that location. Participants then saw a photograph of the location and were asked to assess the degree of correspondence between their thoughts and the photograph. It was predicted that the overall level of perceived correspondence would be high (hypothesis 1), and that participants who believed in the existence of psychic ability would report significantly higher levels of correspondence than others (hypothesis 2).

Proponents of psychic ability could argue that these predicted effects could be due to participants possessing psychic ability, with those who believe in such abilities being especially gifted. A second part of the study assessed this notion by running a series of trials employing blind judging procedures. During each of four trials an experimenter travelled to a distant location and then asked participants to tweet their thoughts about the nature of that location. On each trial, participants then saw five images of different locations (the actual target and four decoys) and were asked to choose the one that best matched their thoughts. This modified procedure controls for possible confirmation bias (see Milton and Wiseman, 1997) and thus any significant effects would support the existence of psychic ability (hypothesis 3). This second stage of the study also acted as a further test of the remote viewing hypothesis, examining the performance of participants who might be expected to produce especially high scores – those who believed in the existence of psychic abilities (hypothesis 4) and those who believed that they possessed psychic abilities and were especially confident about their responses (hypothesis 5), as assessed by an online questionnaire.

While mass participation studies have the benefit of enabling large numbers of individuals to participate in ESP testing, this method can be undermined by an artefact called the 'stacking effect', which occurs when multiple calls are made on a single target (Greville, 1944). Although target guesses may be non-independent (e.g., due to stereotypical thinking or

knowledge of others' guesses), statistical analyses typically assume independence. The simplest solution to this is to use the 'majority vote' technique (Thouless & Brier, 1970), which combines participants' calls for a single target. This technique has had some success in ESP testing (e.g., Barker, Messer & Drucker, 1976; Brier & Tyminski, 1970; Carpenter 1991); indeed it may have a theoretical advantage that goes beyond avoiding a statistical artefact. If there is a genuine but weak effect (in this case, a putative psi signal), it may be detected if calls are combined so that the noise cancels out – a phenomenon that has in mainstream research been termed 'the wisdom of crowds' (Surowiecki, 2005). The present study therefore uses Twitter to engage large numbers of participants and present them with immediate target feedback, and averages their calls to exploit any group-level remote viewing effect. Only five trials were used in order to prove the principle that Twitter could be used to conduct such a study, and because Twitter is typified by short-term interactions.

Procedure

Participants were initially recruited via media reports, resulting in approximately 3000 people following the Twitter account being used to run the experiment. News of the experiment then spread on Twitter, with people sending messages about the study to their networks, resulting in an additional 4000 individuals following the account. Different numbers of participants were involved in each trial, and the method cannot establish which participants took part in all five trials. Each trial took place at 3pm BST on five consecutive days. One hour before the start of each trial the experimenters sent a tweet reminding participants about the forthcoming trial.

Non-blind trial

During the non-blind trial the experimenter traveled to the target location (a weir) and sent a tweet asking participants to submit their thoughts and impressions about the location via twitter. Twenty minutes later the experimenter sent a second tweet asking participants to complete an online questionnaire. This questionnaire asked participants to indicate their sex (response options: 'male', 'female'), whether they believed that psychic ability exists ('definitely yes', 'probably yes', 'uncertain', 'probably no', 'definitely no') and the degree of correspondence between their thoughts and a photograph of the weir ('very high', 'high', 'medium', 'low', 'very low').

Thirty minutes after that the questionnaire was closed.

Blind judging trials

The 20 locations used during the blind judging trials were chosen and photographed prior to the experiment. The images were sorted into 4 sets of 5 photographs, with each set consisting of locations that were visually maximally different. Each image in each set was then randomly labeled A through E, and each set randomly was assigned to one of the four trials.

During each of the four trials an experimenter randomly selected a target from the appropriate set of five possible locations. This involved visiting www.random.org - a website that generates true random numbers via atmospheric noise – and generating a random number between 1 and 5 (where 1= Image A, 2= Image B, etc.). The experimenter then traveled to the chosen location. Once there, they sent a tweet asking participants to submit their thoughts and impressions about the location via twitter. All of the tweets sent after the target had been selected contained a standard wording to ensure that they could not contain any clues about the location.

Twenty minutes later the experimenter then sent a second tweet asking participants to complete an online questionnaire. This questionnaire asked participants to indicate their sex, whether they believed that psychic ability exists, and whether they thought that they had psychic ability ('definitely yes', 'probably yes', 'uncertain', 'probably no', 'definitely no'). Participants were then presented with five photographs, and asked to select the location that best matched their thoughts and rate how confident they were about their guess ('very confident', 'fairly confident', 'not very confident', 'not at all confident').

Thirty minutes after that the questionnaire was closed and a final tweet directed participants to a website revealing the target photograph.

Results

Non-blind trial

As predicted by hypothesis 1, many participants reported a large degree of correspondence between their thoughts and the target, with around 38% of the group rating the correspondence as 'medium', 'high', or 'very high'. As predicted by hypothesis 2, there was a significant relationship between belief in psychic ability and level of perceived correspondence (N=1093, Chi-Square=119.24, DF=16, $p < .0001$, effect size $\Phi = .33$), with those expressing a belief in the paranormal reporting higher levels of correspondence than others (see Table 1).

		Belief in psychic ability					Total
		D ef. no	Prob. no	Uncertain	Prob. yes	Def. yes	
Level of perceived correspondence	Very low	53.39	44.55	31.89	25.38	18.40	36.96
	Low	20.76	26.60	26.49	22.84	20.25	23.70
	Medium	13.56	17.95	23.24	22.84	30.67	20.68
	High	8.47	8.65	12.43	19.80	14.11	12.08
	Very high	3.81	2.24	5.95	9.14	16.56	6.59

Table 1: Percentage of respondents as a function of belief in psychic ability and level of perceived correspondence between their thoughts and the target.

Blind judging trials

Each participant was able to see all of the ‘tweets’ sent by other participants, and thus their responses cannot be considered independent. Because of this, the data was analysed using a ‘majority vote’ technique that treats the group as a single unit of analysis (Brier & Tyminski, 1970). The location that received the greatest percentage of votes on each trial was seen as the group’s selection. If this location matched the actual target location then the trial was deemed a hit. If not, it was deemed a miss. For the experiment to be considered significant, it had to yield three or more hits (associated binomial $p=.02$). On each trial the group failed to identify the correct target, thus hypothesis 3 was not supported.

Participants reporting a strong belief in the existence of psychic ability (answering ‘definitely yes’ to the belief question) identified the correct target on one trial (exact binomial $p=.41$), therefore hypothesis 4 was not supported. Post hoc analysis revealed that there was no significant relationship between belief in psychic ability and choice of target on any of the trials (Trial 1: $N=1322$, Chi-Square= 15.36 , $DF=16$, $p=.50$; Trial 2; $N=1053$, Chi-Square= 13.60 , $p=.62$; Trial 3: $N=815$, Chi-Square= 11.18 , $p=.80$; Trial 4: $N=726$, Chi-Square= 15.11 , $p=.52$).

Finally, those participants who reported that they believed they were psychic (answering ‘definitely yes’ to the psychic ability question) and were confident of their response (answering ‘very confident’ or ‘fairly confident’ to the confidence question) also failed to identify the correct target on any trial, thus providing no support for hypothesis 5.

The target locations, and the locations chosen by the majority of participants, are shown in Table 2.

	Target location	Chosen location		
		All participants	Believe in psychic ability	Believe that psychic and confident
Trial 1	Modern building	Wooded area (N=1322, 35.32%)	Wooded area (N=222, 39.19%)	Wooded area (N=57, 40.35%)
Trial 2	Playpark	Set of stairs (N=1053, 26.69%)	Playpark (N=183, 29.51%)	Set of stairs (N=50, 42.00%)
Trial 3	Unusual canopy	Graveyard (N=815, 24.05%)	Split vote between tunnel and bridge (N=143, 23.08%)	Tunnel (N=37, 29.73%)
Trial 4	Post box	Canal (N=726, 23.83%)	Canal (N=127, 26.77%)	Canal (N=35, 34.29%)

Table 2: Targets selected by all participants, those who believed in psychic ability, and those who believed they possessed psychic ability and were confident about their responses (total number of participants and percentage who chose the location in parentheses).

Discussion

In the non-blind trial the overall level of perceived correspondence was high, and participants who believed in the existence of psychic ability reported significantly higher levels than others. In the blind trials (which eliminated potential confirmation biases) participants were unable to

accurately identify the target locations, and there was no significant difference between the performance of believers and others. In addition, the one group of participants that might be expected to produce especially high scores – those who reported possessing psychic abilities and who were especially confident about their responses – failed to identify any of the targets. Participants who believed in the existence of psychic ability identified one target, which is close to chance performance. These findings do not support the existence of remote viewing. They do, however, suggest that the use of non-blind judging procedures in remote viewing research is likely to inflate the level of perceived correspondence between participants' thoughts and a target location, especially for those who believe in the existence of such abilities. Interestingly, such uncontrolled procedures are frequently used by those running public workshops and seminars that claim to develop individuals' remote viewing skills, and thus could easily convince participants that they do indeed possess such abilities. In addition, although the present study explored this effect within the context of a remote viewing study, it is easy to imagine how the same mechanism could cause people to believe that their dreams predict the future or that an astrological reading accurately describes their past. As such, an individual's ability to find illusory correspondences between their thoughts and external events could play a key role in the formation and maintenance of belief in the paranormal.

The study does have some limitations. While it attempted to tap the wisdom of the crowds by using a majority vote method, it was not able to detect any evidence for remote viewing. With only four blind trials, though, there is a risk of committing a Type II error (failing to detect a genuine effect) and we would recommend using a greater number of trials in future research using this method (we could not estimate statistical power beforehand due to the difficulty in accurately estimating effect size with an untested method that also uses majority vote techniques.) Additionally, the majority vote method will have the effect of detecting any group-level effect, and this could include group biases (e.g, a stereotypical preference for the leftmost target in an array of target possibilities; ideally the order of presentation of the array of target possibilities should be randomised so as to eliminate this potential bias). This kind of group level effect could outweigh any more subtle group psi effect, once again suggesting that a larger number of trials would be needed in future.

This experiment also examined the potential of Twitter as a research tool, using this relatively new form of social media to both recruit and run participants. The results were, for the most part, very encouraging. Several thousand people signed up for the experiment and each trial

attracted around a thousand people who contributed their data in a very short space of time. Although there was some attrition from one trial to the next, the study maintained a healthy rate of participation. In addition, the study attracted participants with a wide range of beliefs about psychic ability, and the large numbers involved meant that even focused sub-groups of individuals, such as those believing that they were psychic and confident about their responses, contained healthy numbers of people.

The remote viewing study worked well because it caught the interest of both the media and public alike, and utilised Twitter's ability to provide a near instant response from participants, and to give them near instant feedback. It seems likely that other studies that share these features are also likely to succeed on Twitter. This could include, for example, studies using the 'experience-sampling method', wherein participants report their thoughts and feelings the moment they receive a certain signal (see, e.g., Larson & Csikszentmihalyi, 1983; Hurlburt & Heavey, 2004). The inter-connectedness of Twitter users, combined with the potential for messages about studies to reach a large audience, also means that it may also prove a highly effective way of recruiting specialised groups of participants.

In short, this study did not yield any evidence for remote viewing, but did show that the perceived correspondences between a participant's thoughts and target in uncontrolled remote viewing trials could be due to confirmation bias, and that individuals who believe in the paranormal are especially prone to this bias. The study also showed that Twitter can be used as a new form of research tool, and can be an effective way of recruiting and running participants. It seems that the nature of Twitter makes it especially well suited for studies that require an immediate response from a large number of participants, and that would benefit from giving participants immediate feedback. We look forward to other experimenters utilising this tool in creative ways.

Acknowledgements

We are grateful to all the members of the public who participated in this study, and to our referees for their helpful comments.

References

Barker, P., Messer, E., and Drucker, S. (1976). A group majority vote procedure with receiver optimisation. In J. D. Morris, W. G. Roll & R. L. Morris (Eds.), *Research in Parapsychology 1975*. Metuchen, N.J: Scarecrow Press.

Blackmore, S. J. (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.

Brier, R., and Tyminski, R. (1970) Psi application: Part II. The majority-vote technique – Analyses and observations. *Journal of Parapsychology*, 34, 26-36.

Brugger, P., and Graves, R. E. (1997). Testing vs. believing hypotheses: Magical ideation in the judgment of contingencies. *Cognitive Neuropsychiatry*, 2, 55-57.

Brugger, P., Regard, M., Landis, T., Cook, N., Krebs, D., and Niederberger, J. (1993). “Meaningful” patterns in visual noise: Effects of lateral stimulation and the observer's belief in ESP. *Psychopathology*, 26, 261-265.

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

Carpenter, J. C. (1991). Prediction of forced-choice ESP performance: Part III. Three attempts to retrieve coded information using mood reports and a repeated-guessing technique. *Journal of Parapsychology*, 55, 227-280.

French, C. C., Herrmann, D., Hales, S., and Northam, C. (1996) Psi in the eye of the beholder? Sheep-goat differences in judging paranormal evidence. XXVI International Congress of Psychology, Montreal, Canada, 16-21 August. (Abstract published in the *International Journal of Psychology*, 31, 260.)

Goodfellow, L. D. (1938). A psychological interpretation of the results of the Zenith radio experiments in telepathy. *Journal of Experimental Psychology*, 23, 601-632.

Greville, T. N. E. (1944). On multiple matching with one variable deck. *Annals of Mathematical Statistics*, 15, 432-434.

Hurlburt, R., and Heavey, C.L. (2004). To Beep or Not To Beep: Obtaining Accurate Reports About Awareness. *Journal of Consciousness Studies*, 11(7-8), 113-128.

Hyman, R. (1995). Evaluation of the program on anomalous mental phenomena. *Journal of Parapsychology*, 59(4), 321-351.

Larson, R., and Csikszentmihalyi, M. (1983). The experience sampling method. *New Directions for Methodology of Social and Behavioral Science*, 15, 41-56.

Marks, D. (2000). *The psychology of the psychic (2nd edn)*. New York: Prometheus Books.

Milton, J., and Wiseman, R. (1997). *Guidelines for extrasensory perception research*. Hatfield, UK: University of Hertfordshire Press.

Milton, J., and Wiseman, R. (1999). A meta-analysis of mass media ESP testing. *British Journal of Psychology*, 90, 235-240.

Mumford, M. D., Rose, A. M., and Goslin, D. A. (1995). *An evaluation of remote viewing: Research and applications. The American Institutes for Research report*. September 29.

Pizzagalli, D., Lehmann, D., and Brugger, P. (2001). Lateralized direct and indirect semantic priming effects in subjects with paranormal experiences and beliefs. *Psychopathology*, 34, 75-80.

Surowiecki, J. (2005). *The wisdom of crowds: Why the many are smarter than the few*. London: Abacus.

Thouless, R. H., and Brier, R. M. (1970). The stacking effect and methods of correcting for it,

Journal of Parapsychology, 34, 124-128.

Utts, J. (1995). An assessment of the evidence for psychic functioning. *Journal of Parapsychology*, 59(4), 289-320.

Wiseman, R., and Smith, M. D. (2002). Assessing the role of cognitive and motivational biases in belief in the paranormal. *Journal of the Society for Psychical Research*, 66, 157-166.