APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **16**: 17–34 (2002)
DOI: 10.1002/acp.754

The Connection Between Random Sequences, Everyday Coincidences, and Belief in the Paranormal

PAOLA BRESSAN*

Università di Padova, Italy

SUMMARY

This paper argues against the theory that people interpret unusual coincidences as paranormal because they misunderstand the probability of their occurring by chance. In the two studies reported here, 214 subjects were given a questionnaire on the frequency of coincidences in their lives, a series of probabilistic problems, and a scale assessing their belief in the paranormal. Believers reported more coincidences than disbelievers. Believers made more errors than disbelievers in tasks reflecting sensitivity to the relationship between expected distribution of chance events and total number of occurrences; and avoided repetitions of identical alternatives in a random sequence to a greater extent. However, the last two effects completely disappeared in a subsample of university students. It is proposed that a more frequent experience of coincidences, on the one hand, and a more biased representation of randomness, on the other, are independent consequences of a stronger propensity of believers in the paranormal to connect separate events. Copyright © 2002 John Wiley & Sons, Ltd.

Many people believe in the paranormal (see Gallup and Newport, 1991). The most frequent reason given for believing is a personal paranormal experience, like thinking of something at the exact moment at which the person one is with starts to say it, or having a premonition or dream that then comes true (Clarke, 1995). Yet phenomena such as telepathy or precognition have never been reliably demonstrated in controlled conditions: many anecdotal cases and few experimental results, none of which has proved consistently replicable, represent the sole evidence for their existence. This is for some an attestation of their predictably elusive nature, but for others evidence that paranormal phenomena do not exist at all, and that the paranormal flavour of some experiences must be a matter of mistaken interpretation, the consequence of a cognitive bias.

In particular, it has been suggested (Blackmore and Troscianko, 1985; Brugger *et al.*, 1994) that the preference for a paranormal explanation in the face of an anomalous event may result from a lack of appreciation of how often that particular event will occur by chance. If this probability is underestimated, the event will, as all non-chance events, call for a cause – a paranormal cause if natural ones do not fit.

The empirical evidence for the relationship between probabilistic errors and paranormal belief, however, is mixed. Believers and disbelievers (picturesquely referred to in the

*Correspondence to: Paola Bressan, Dipartimento di Psicologia Generale, Università di Padova, Via Venezia 8, 35131 Padova, Italy. E-mail: paola.bressan@unipd.it

literature as sheep and goats, after Schmeidler, 1943) do not seem to differ in either intellectual ability, as measured by the Raven's Advanced Progressive Matrices test, or statistical competence, as measured by their grade on a statistics exam (Thalbourne and Nofi, 1997). Yet a slight association between paranormal belief and faulty conditional reasoning has been observed (Wierzbicki, 1984, not replicated by Irwin, 1991; Roberts and Seager, 1999). Likewise, Blackmore and Troscianko (1985) did find that goats tended to perform better than sheep in some types of problems requiring judgements of probability, though this result (not replicated by Roberts and Seager, 1999) should be viewed with caution, because the problems to which it applied were not the same in a sample of high-school female students and in a sample of university students, and the (just) significant differences obtained in one group were not observed in the other.

The type of probabilistic errors of true interest here should be those that concern the internal representation of chance. It is well known that humans are unable to produce a random series of discrete responses, even when instructed to do so (see Bar-Hillel and Wagenaar, 1993, for a review). Subjectively 'natural' random sequences contain an excess of alternations between different elements, and too much balancing of such elements over relatively short segments. Do believers in the paranormal manifest a stronger bias than disbelievers?

The answer to this question is 'perhaps'. In a study by Brugger *et al.* (1990), sheep, more than goats, avoided repetitions when trying to produce random strings of the digits 1–6, and considered, when presented with pairs of equiprobable short random sequences, more likely the sequence of the pair that contained fewer repetitions. When asked to produce long random strings of four alternatives (rabbit, duck, carrot, and marsh reed), sheep, more than goats, avoided consecutive repetitions of those semantically related, such as rabbit and carrot (Brugger *et al.*, 1995). Yet no effect of belief on the number of repetitions in random-generation tasks was found by Blackmore and Troscianko (1985), or by Blackmore *et al.* (1994).

People who underestimate the probability that two identical digits occur one after another by chance may, in everyday life, underestimate the probability that two similar events occur one after another by chance. In this case, stronger underestimators should experience subjectively non-random ('meaningful') coincidences more frequently than weaker underestimators. If sheep have a more biased representation of chance than goats, then they should report coming across more coincidences. The evidence for this argument is, again, mixed. On the one hand, Brugger *et al.* (1990) did find a moderate correlation (Spearman $rho = 0.27$) between belief in the paranormal and subjective frequency of coincidences. On the other, Blackmore (1997) found that sheep and goats did not differ in their assessment of the probability of observing a certain coincidence in the general population; and neither sheep nor goats underestimated it.

The work reported in this paper has some bearing on two related issues. The first stems from the empirical disagreement, and is handled through three proofs of existence performed together on a large number of people: do believers in the paranormal tend to make more errors in probabilistic reasoning? Do they have a more biased representation of randomness, as manifested in repetition avoidance? Do they come across more coincidences? The second, and foremost, objective of this study was that of carrying out a critical test of the concept that the frequency of coincidences ensues from a misrepresentation of chance. Is there a correlation between frequency of coincidences in everyday life and repetition avoidance?

STUDY 1

Method

Participants

The subjects were 111 (38 men and 73 women), ranging in age from 19 to 62 years. They were recruited in public places and examined individually. Seventy subjects (22 males and 48 females) were university students of various faculties: Psychology, Letters, Engineering, Biology, and Philosophy. Forty-one subjects (16 males and 25 females) were regular workers: mostly unskilled and skilled workers, clerks, salespersons. About half of them had obtained a middle-school certificate, the other half a high-school certificate. Median ages were 21 years for the university student subsample, 32 years for the worker subsample.

Procedure

Each participant was asked to complete three questionnaires, in the following order: a questionnaire on coincidences, a questionnaire containing a number of problems requiring probabilistic reasoning, and a scale on paranormal belief. The paranormal belief scale was presented last so that responses to the coincidence and reasoning questionnaires could not possibly be biased by subjects knowing that the study was concerned with paranormal belief. For this reason, care was also taken to avoid words with direct reference to the paranormal (such as ‘telepathy’ or ‘precognition’) in the coincidence questionnaire. All questionnaires were in Italian. There was no time limit; completion of the test took from 30 to 45 minutes.

Materials

Coincidence questionnaire

The questionnaire on coincidences was constructed keeping in mind the measuring instruments and results of a large-scale UK survey on coincidence experiences (Henry, 1993). It concerned the frequency with which subjects had experienced certain categories of coincidences; the inventory separately assessed first the frequency of coincidences in general, and then the frequency of coincidences of different categories. At the end, the questionnaire listed a number of possible causes of coincidences, and subjects had to indicate whether they believed in such causes. The complete questionnaire is shown in Appendix A.

Probabilistic reasoning questionnaire

The probabilistic reasoning questionnaire consisted of five problems, concerning the comprehension of sampling issues, the sensitivity to sample size, and the generation of random sequences. The order of problems was randomized across subjects. The questionnaire is described in Appendix B, together with the methods used for scoring.

Belief in the paranormal questionnaire

The belief in the paranormal was assessed through a specially constructed scale consisting of 20 items. These were statements concerning telepathy, precognition, clairvoyance, psychokinesis, divination, spiritualism, and witchcraft (sample items for each category: ‘The direct communication between two minds is possible’; ‘Certain premonitions and presentiments come true in such detail that they cannot be due to chance'; 'Some psychics can "see" events distant in space, for example indicate the place where a person who has disappeared can be found'; 'There are gifted individuals who can move or bend certain objects through mental forces only'; 'Through card reading, a good fortune-teller is able to predict future events'; 'During a seance, it is possible to contact the spirits of the dead'; 'Amulets keep negative forces away'). The response to each item was made on a 7-point Likert scale, ranging from 1, corresponding to 'I do not believe it', through a midpoint of 4, 'I don't know', to 7, corresponding to 'I believe it'. The paranormal belief score was the mean of the ratings of the individual items. The scale has been tested and found to have good internal consistency (Cronbach's alpha = 0.93) and test–retest reliability (rtt = 0.87, $p < 0.0001$, over 6 weeks, for a 28-subject sample), as well as construct validity; its factor structure and psychometric properties will be described elsewhere (Bressan, in preparation).

Results

Data were analysed via Spearman's rank-order correlations and Student's t-tests. All significance levels reported throughout the paper are two-tailed.

Paranormal belief

Scores were normally distributed, with a mean of 3.41 (SD 1.27) and range 1.15–6.55; the mean was higher for workers (3.9) than for students (3.1), $t = 2.97$, $p = 0.004$. Subjects with scores of 2.45 (corresponding to the 25th percentile) or less were classified as goats, those with scores over 4.15 (corresponding to the 75th percentile) as sheep; the remaining subjects were classified as intermediates. There were 28 goats, 28 sheep and 55 intermediates.

Probabilistic reasoning

The mean number of correct responses in the four-part sampling task was 2.08 (SD 0.56); there was no significant difference between the performances of goats (mean 2.2) and sheep (mean 2.0), $t < 1.5$, n.s. (This is in agreement with the data reported by Roberts and Seager, 1999.) As to the maternity ward problem, 21% of subjects thought that the larger hospital was more likely than the smaller hospital to record days on which the babies born were in majority boys, and 68% that the probability was the same. These proportions substantially correspond to those reported by Kahneman and Tversky (1972). Again, performance was not associated with paranormal belief. Only 13 subjects (12%) gave the correct answer: of these, six were goats and four sheep.

Generation of randomness

Spearman's rank-order correlation between paranormal belief scores and the number of repetitions was negative and significant in all three random-generation tasks. The association was quite small both in Blackmore and Troscianko's replication task ($rho = -0.19$, $p = 0.047$) and in the hidden-string task ($rho = -0.20$, $p = 0.035$), and increased in the visible-string task ($rho = -0.33$, $p < 0.0001$). When the data were separately analysed for students and non-students, it was found that all correlations only held within the latter group. The correlation between paranormal belief score and number of repetitions in the most discriminative task (visible string), for instance, was -0.66 ($p < 0.0001$) in the worker subsample, and zero (-0.08, n.s.) in the student subsample.

Self-reported frequency of coincidences
The frequencies reported in the seven categories of coincidences were added to give a combined coincidence frequency score. No differences between students and non-students emerged in the frequency of coincidences, either general or combined (as shown by Mann–Whitney tests).

There was no correlation whatsoever between frequency of coincidences, either general or combined or of a particular class, and performance in any of the five problems. However, a remarkable positive correlation emerged between paranormal belief score and both general ($rho = 0.46$, $p < 0.0001$) and combined ($rho = 0.36$, $p = 0.0002$) frequency of coincidences. Such correlations were positive and significant in both the worker and student subsamples. The median combined frequency score for coincidences was 16 for goats, 20 for intermediates, and 23 for sheep (Kruskal-Wallis $H = 18.3$, $p < 0.0001$).

Paranormal believers, then, did report coming across more coincidences; this was a large effect and was independent of the educational level. They did not make more errors in tasks requiring probabilistic reasoning, at least in the form of sensitivity to the effects of replacement (or lack thereof) in sampling, but showed a more pronounced avoidance of repetitions in random sequences. When workers and university students were separately considered, however, it was found that this association was large in the first group, and completely disappeared in the second.

The second question of interest was, does the higher subjective frequency of coincidences in believers follow from their more biased representation of chance, as expressed in repetition avoidance? The data suggest it does not: frequency of coincidences in everyday life and repetition avoidance were not associated. This finding is especially surprising in the worker subsample, where both variables correlate with paranormal belief.

The tangible difference between students and non-students in the association between paranormal belief and repetition avoidance was a new and unexpected finding, also considering that previous evidence both for and against the existence of such association has always been collected in samples of students. In the light of the existing speculations on the causes of belief, the lack of association between coincidence frequency and repetition avoidance was a surprising result, too. It was then decided to determine if these findings could withstand replication in a separate sample.

STUDY 2

Method

Participants
A new sample of 103 subjects (53 men and 50 women), ranging in age from 19 to 61 years, took part in the study. They were recruited in public places and examined individually. Forty-five subjects (22 males and 23 females) were university students of various faculties: Psychology, Engineering, Letters, Economics, Philosophy, Biology, and Architecture. Fifty-eight subjects (31 males and 27 females) were regular workers; about half of them had obtained a middle-school certificate, the other half a high-school certificate. Median ages were 23 years for the university student subsample, 29 years for the worker subsample.

Procedure
The procedure was the same as in Study 1; completion of the test took from 30 to 60 minutes.

Materials
The coincidence and the paranormal belief questionnaires were identical to those of Study 1. The probabilistic reasoning questionnaire was expanded to include eight problems, probing into the representativeness bias (as applied to sample size or with size kept constant), and the generation of randomness. The order of problems was randomized across subjects. The questionnaire is described in Appendix C, together with the methods used for scoring.

Results

Paranormal belief
Scores were normally distributed, with a mean of 3.35 (SD 1.41) and range 1.00–6.60. Subjects with scores of 2.10 (corresponding to the 25th percentile) or less were classified as goats, those with scores over 4.40 (corresponding to the 75th percentile) as sheep; the remaining subjects were classified as intermediates. There were 26 goats, 28 sheep and 49 intermediates.

As in Study 1, the mean paranormal belief score was higher for workers (mean = 3.8) than for students (mean = 2.7), $t = 4.35$, $p < 0.0001$.

Representativeness bias as applied to sample size
In the pollster problem, only 35% of respondents answered correctly, that is, were aware of the fact that larger samples provide more reliable estimates. Thirteen per cent favoured the smaller sample, whereas 52% thought that one should have equal confidence in the two estimates. These results indicate an even stronger bias than that found by Bar-Hillel (1982), who explains it with the fact that a sample's representativeness is determined not only by the sample size but by the ratio between the sample size and the population size as well. In this case, these ratios are the same (1 out of every 1000), thus the two samples tend to be regarded as equally representative. The point-biserial correlation between scores (0 or 1) and paranormal belief was zero ($r = -0.07$, n.s.).

Students performed significantly better than workers (mean scores were 0.42 versus 0.07, $t = 2.8$, $p = 0.006$). The point-biserial correlation with paranormal belief was non-significant in both the worker ($r = -0.04$, n.s.), and student ($r = 0.14$, n.s.) subsamples.

In the coin-tossing problem, about half of the subjects scored zero, that is, were unaware of the fact that the likelihood of obtaining a certain distribution of outcomes depends on the number of tosses. Means were −0.17 for the 50% proportion, 0.17 for the 75% proportion, and 0.38 for the 100% proportion; all differences are significant at least at the 0.002 level (paired-samples t's). The task was easiest for the most extreme deviation (100% heads) from the proportion of heads expected by chance: in this case, 41% of subjects could understand that the larger the number of tosses, the less likely it became that a fair coin would always be falling with its 'head' face up. Twenty-eight per cent understood this for the 75% proportion of heads. These data support Bar-Hillel's (1982) contention that the more a sample result deviates from the result expected in its parent population and is thus perceived as non-representative, the easier it is to grasp that it will be less likely in a larger sample. Only 9% of subjects saw that when the proportion of heads is 50%, larger numbers of tosses make it more likely that the coin is unbiased.

A significant correlation emerged between paranormal belief and total score ($rho = -0.38$, $p < 0.0001$): mean scores were 1 for goats and -0.1 for sheep, $t = 4.1$, $p < 0.0001$; intermediates were in between with a score of 0.3. The difference between goats and sheep was smallest for the 100% problem (means were 0.46 versus 0.29, $t = 1.16$, n.s.), increased in the 75% problem (0.38 versus 0, $t = 2.39$, $p = 0.02$), and was largest for the 50% problem (0.15 versus -0.39, $t = 3.89$, $p < 0.0001$). Believers, more than disbelievers, failed to notice that a perfect 50:50 proportion is more representative of a larger, rather than smaller, sample: the underlying idea seems to be that 200 heads out of 400 tosses decrease the likelihood that the coin is fair compared to two heads out of four, because chance is imperfect and in the long run coins are doomed to produce a few minor irregularities.

Students performed significantly better than workers in this task (mean scores were 0.76 versus 0.07, $t = 3.5$, $p < 0.001$). The correlation with paranormal belief was separately checked in each group, and turned out to be significant ($rho = -0.36$, $p = 0.006$) in the worker subsample, but virtually zero ($rho = -0.01$, n.s.) in the student subsample.

Representativeness bias as applied to random sequences

In the birth-order problem, the correct answer was 'about 70', but the mean estimate was 35. Over 80% of the subjects gave an estimate smaller than 70, and only less than 4% of subjects chose a larger number: we are therefore facing here an extremely robust cognitive illusion. These results are in agreement with those of Kahneman and Tversky (1972), who interpret the error as a consequence of the fact that the first sequence appears more representative than the second (which does not reflect the equal proportion of males and females in the population). No significant association emerged between the performance in this task and paranormal belief. The point-biserial correlation between scores (0 or 1) and paranormal belief was zero ($r = -0.05$, n.s.). There were no significant performance differences between students and workers.

In the dice-rolls problem, the mean number of correct responses (out of 6) was 2.74. Only 22% of the subjects performed the whole task correctly, whereas 60% scored 2 or less. There was a moderate negative correlation between number of correct responses and paranormal belief ($rho = -0.27$, $p = 0.006$): sheep (mean = 2.0) answered less correctly than goats (mean = 3.7), $t(52) = 2.79$, $p = 0.007$. The performance of the intermediate subjects was in between (mean = 2.6). These results represent a nearly perfect replication of the findings by Brugger *et al.* (1990).

The performance difference between students (mean = 3) and workers (mean = 2.5) was non-significant ($t < 1.2$). The correlation between performance and paranormal belief held, with about the same strength, for both students ($rho = -0.27$) and workers ($rho = -0.28$).

Generation of randomness

The correlation between paranormal belief and number of repetitions was negative and significant in the six-alternative random-generation task, both in the hidden ($rho = -0.26$, $p = 0.007$) and visible ($rho = -0.36$, $p < 0.0001$) conditions, which replicated the effect found in Study 1. The association with paranormal belief was only significant in the worker subsample. In the most discriminative task (visible string), correlations were

-0.48 ($p < 0.0001$) in the worker subsample, and zero (-0.08, n.s.) in the student subsample. This, again, replicated the difference found in Study 1.

In the two-alternative (Head or Tails) task, the mean number of consecutive repetitions of the same alternative (mean $= 26.3$) was significantly smaller than the number of repetitions mathematically expected in a sequence of 66 coin tosses, which is 32.5 (one-sample $t = -7.6$, $p < 0.0001$). There were no significant performance differences between students and workers. The correlation between paranormal belief score and number of repetitions did not reach statistical significance, although the means were in the right direction (24.6 for sheep, 26.4 for intermediates, 27.8 for goats).

Self-reported frequency of coincidences
No significant correlation emerged between frequency of coincidences, either general or of a particular class, and scoring in any of the problems.

There was a clear positive correlation between paranormal belief score and both general ($rho = 0.51$, $p < 0.0001$) and combined ($rho = 0.66$, $p < 0.0001$) frequency of coincidences, which replicated the findings of Study 1. Both correlations had about the same magnitude for students and workers. The median combined frequency score for coincidences was 17 for goats, 20 for intermediates, and 24 for sheep (Kruskal–Wallis $H = 36.4$, $p < 0.0001$).

STUDIES 1 AND 2: COMBINED RESULTS AND DISCUSSION

Probabilistic reasoning

There were no performance differences between sheep and goats in the four-part sampling task, the pollster, maternity ward, and birth-order problems; but scoring in the coin tossing task was indeed negatively and significantly associated with paranormal belief, as was repetition avoidance in the dice-rolls problem. The pattern of these data suggests that the reason why no difference between believers and disbelievers emerged in some of the tasks is, simply, that they were too difficult.

All these tasks call into play the representativeness heuristic (Kahneman and Tversky, 1972; Bar-Hillel, 1982), i.e. one of the cognitive shortcuts to which people resort when considering the probability of some uncertain event. An event's representativeness is given by the extent to which the event itself appears to reflect the salient features of its parent population, or, alternatively, of the process by which it is generated. If the event in question is a proportion expected by chance (such as in the maternity ward and coin-tossing tasks) or a random sequence (such as in the birth-order and dice-rolls problems), its subjective likelihood will depend on its capability to reflect the fundamental characteristics of randomness, which are mainly lack of patterns and balanced occurrence of all alternatives. In this sense, the pollster problem is different from the others, in that it probes into the nature of the representativeness heuristic (a representative result is equally likely in a large or small sample, when sample-to-population ratios are the same), but does not call into play the psychological attributes of randomness. The lack of differences between sheep and goats in this problem may well reflect a lack of differences in the essential features of the representativeness bias, more specifically a disregard of absolute, in favour of relative, sample size.

In all other tasks, the use of the representativeness heuristic by our respondents is only the machinery through which we come to discover something on their representation of randomness. The maternity ward and the coin-tossing problems reflect the same question: is a non-representative result (notably, a deviation from the expected equipresence of alternatives, that is, from a 50:50 proportion of either Boys and Girls or Heads and Tails) more likely in a large or a small sample? It has been shown that a correct solution is much easier for more extreme deviations. In the maternity ward problem the observed proportion is 60:40, a deviation that is not large enough to activate the 'non-representative' label (as argued by Bar-Hillel, 1982). In this case, the bias is compelling and virtually universal: only 12% of subjects responded correctly, and there were no performance differences between sheep and goats. Consistent with Bar-Hillel's account, respectively 28% and 41% of subjects gave the correct answer when, in the coin-tossing problem, the observed proportions became 75:25 and 100:0. Here, a less markedly skewed distribution of responses allowed the emerging of a difference between sheep and goats.

In the two tasks discussed above, the stress is on the relationship between non-representative distributions and number of outcomes. In the birth-order and dice-rolls problems, the same issue is looked at from another perspective, and the emphasis is on how strongly non-representative a distribution appears, when the number of outcomes is kept constant (sequence length is 6 in either problem). In both cases, two sequences are compared and one contains more consecutive repetitions than the other. In the birth-order problem, the possible alternatives are two, Boy and Girl; in the dice-rolls problem, the possible alternatives are six (digits 1–6). Goats did not perform significantly better than sheep in the Boy and Girl task, which corroborates, in a less abstract context, the finding that in random generation there were no significant sheep/goat differences when the alternatives were two (Heads and Tails). Obviously enough, the expression of repetition avoidance is severely limited when the elements to choose from are only two. A larger number of alternatives permits a wider range of inter-individual variations, and, again, a manifesting of differences between believers and disbelievers.

Paranormal belief

The paranormal belief scale, the coincidence frequency questionnaire and the two 6-alternative random-generation tasks were identical in Experiments 1 and 2. Therefore, the corresponding data were combined and analysed together in greater detail.

Paranormal belief scores were normally distributed, with a mean of 3.38 (SD 1.34) and range 1.00–6.60. Subjects with scores of 2.40 (corresponding to the 25th percentile) or less were classified as goats, those with scores over 4.36 (corresponding to the 75th percentile) as sheep; the remaining subjects were classified as intermediates. There were 56 goats, 53 sheep, and 105 intermediates.

Subjects' age had no effect on the degree of their belief ($rho = 0.13$, n.s.). The mean paranormal belief score was higher for workers (mean = 3.89) than for students (mean = 2.97), $t = 5.11$, $p < 0.0001$, and was slightly higher for females (mean = 3.55) than for males (mean = 3.15), $t = 2.15$, $p = 0.03$.

Generation of randomness

The mean number of consecutive repetitions of the same digit was, in both the hidden (mean = 6.7) and the visible (mean = 7.4) conditions, significantly smaller than the mathematically expected number of consecutive repetitions in a sequence of 66 dice rolls, which is 10.8 (both one-sample t's are significant at least at the 0.0001 level). There was a tendency for students to make more repetitions, both in the hidden-string (where the mean number of repetitions was 7.1 for students and 6.1 for workers) and in the visible-string condition (7.9 for students and 6.8 for workers), but neither difference reached significance (both t's < 1.5). There were no differences between females and males (both t's < 1).

Spearman's rank-order correlation between paranormal belief score and number of repetitions was negative and significant in both the hidden-string task ($rho = -0.23$, $p = 0.001$), and the visible-string task ($rho = -0.35$, $p < 0.0001$). The mean number of repetitions is shown in Figure 1, separately for sheep, goats, and intermediates. The effect had the same direction and size for both digrams, i.e. runs of two identical digits, and trigrams, i.e. runs of three identical digits; in the visible-string task, for instance, goats used on average 6.9 digrams and sheep 3.7, $t = 4.7$, $p < 0.0001$; goats used one trigram and sheep 0.19, $t = 4.6$, $p < 0.0001$. For both digrams and trigrams, means were separated by about one standard deviation, a large effect in Cohen's (1988) terminology.

Interestingly, the possibility of seeing the previous responses had opposite effects on goats and sheep: for goats, it made the string more random, i.e. reduced the bias, increasing the number of repetitions by an average of 1.25 (paired-samples $t = -2.6$, $p = 0.012$); for sheep, it made the string less random, i.e. augmented the bias, decreasing the number of repetitions by an average of 0.85 (paired-samples $t = 2.97$, $p = 0.004$). In other words, when allowed to 'adjust' the sequence they were producing according to their representation of chance, disbelievers made it less distorted, and believers more.

Self-reported frequency of coincidences

There were no differences between students and workers in the general (in both cases the median was 3, corresponding to 'a few times'), and combined (in both cases the median

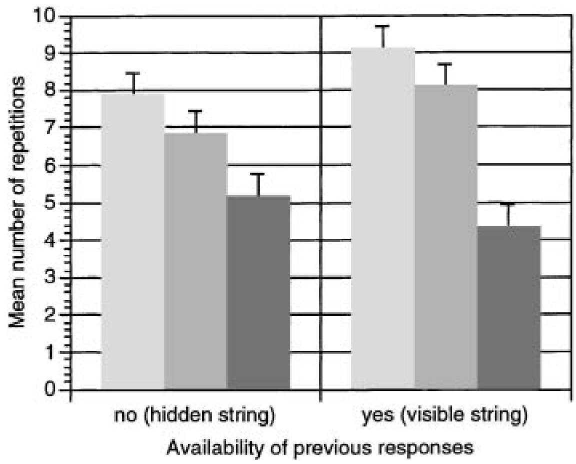

Figure 1. Mean number of repetitions in a random generation task, plotted as a function of the possibility of seeing previous responses (no versus yes), separately for goats (light bars), intermediates (grey bars), and sheep (black bars). The number of repetitions actually expected by chance is 10.8. Error bars indicate the standard error of the mean

was 20) frequency of coincidences, as shown by Mann–Whitney tests. Spontaneous associations and small-world encounters were reported to occur most often, followed in order by perception of something distant in time, cluster experiences, guardian-angel experiences, unexpected solutions of problems, and perception of something distant in space. Spearman correlations between classes of coincidences were all positive and significant, and ranged from 0.13 to 0.41.

It is interesting to note that 75% of subjects believe they have had at least one precognitive experience (perception of something distant in time), and 45% a telepathic or clairvoyant experience (perception of something distant in space). Compared to males, females reported a significant or marginally significant higher rate of coincidences of all categories, but the differences were small. Perhaps curiously, there was no significant correlation between coincidence frequency and age, with the only exception that coincidences interpretable as instances of telepathy or clairvoyance tended to be negatively associated with age among non-students, $rho = -0.22$, $p = 0.03$.

There was no significant correlation between frequency of coincidences, either general or combined or of a particular class, and repetition avoidance. The correlations between paranormal belief score and both general ($rho = 0.49$, $p < 0.0001$) and combined ($rho = 0.52$, $p < 0.0001$) frequency of coincidences were positive and significant. The median combined frequency score was 17 for goats, 20 for intermediates, and 24 for sheep (Kruskal–Wallis $H = 51.0$, $p < 0.0001$).

In Figure 2, this association is plotted as mean coincidence frequency score per category, separately for sheep, goats, and intermediates. The categories of coincidences most remarkably correlated with the paranormal belief score were perception of something distant in time ($rho = 0.48$, $p < 0.0001$), small-world encounters ($rho = 0.39$, $p < 0.0001$), perception of something distant in space ($rho = 0.38$, $p < 0.0001$), and spontaneous associations ($rho = 0.37$, $p < 0.0001$), followed by coincidences of the series or clusters type ($rho = 0.28$, $p < 0.0001$), and unexpected solutions of a problem

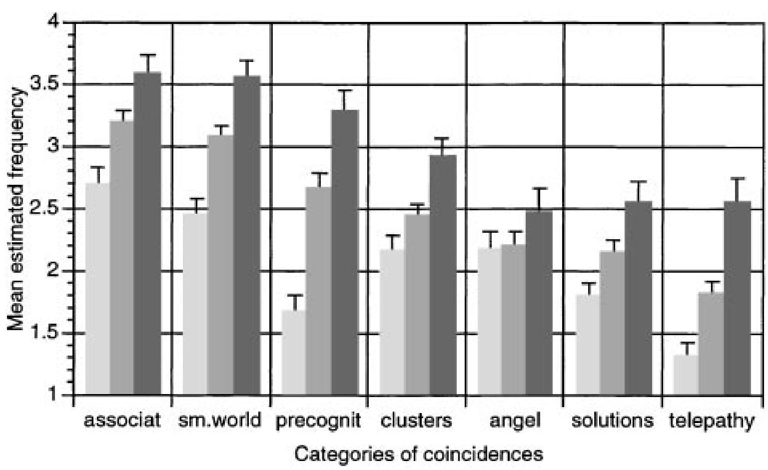

Figure 2. Mean reported frequency of everyday coincidences, plotted as a function of category (spontaneous associations, small-world encounters, perception of something distant in time, series or clusters, guardian-angel experiences, unexpected solutions of a problem, perception of something distant in space), separately for goats (light bars), intermediates (grey bars), and sheep (black bars). Error bars indicate the standard error of the mean

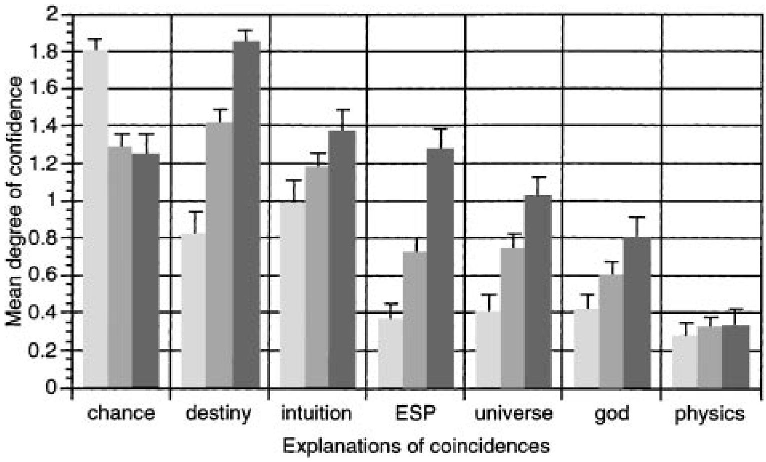

Figure 3. Mean belief in possible explanations of everyday coincidences, plotted as a function of type of explanation (chance, destiny, intuition, extra-sensorial perception, the fact that everything is connected to everything else in the universe, divine intervention, a physical principle not yet discovered by science), separately for goats (light bars), intermediates (grey bars), and sheep (black bars). Error bars indicate the standard error of the mean

($rho = 0.24$, $p < 0.0001$). Only guardian-angel experiences did not correlate at all ($rho = 0.08$, n.s.).

Explanations of coincidences

For each of the proposed causes of coincidences, subjects were given a score of 2 if they chose the answer 'yes', 1 if they chose the answer 'don't know', and 0 if they chose the answer 'no'. Thus, increasing scores indicated increasing confidence in that particular explanation. The only difference between the sexes was in a somewhat higher inclination of females to attribute coincidences to destiny and to the fact that everything is connected to everything else in the universe, but the corresponding effect sizes were small (means were separated by less than 0.3 standard deviations). Overall, the preferred explanations were mere chance, destiny, and intuition; the least frequently chosen was a physical principle not yet discovered by science. Apart from this one, all explanations were significantly correlated with the paranormal belief score, especially destiny ($rho = 0.43$, $p < 0.0001$), extra-sensorial perception ($rho = 0.41$, $p < 0.0001$), and the fact that everything is connected to everything else in the universe ($rho = 0.30$, $p < 0.0001$); more weakly correlated were divine intervention ($rho = 0.22$, $p = 0.002$) and intuition ($rho = 0.17$, $p = 0.011$). The 'mere chance' explanation correlated negatively ($rho = -0.26$, $p < 0.0001$). In Figure 3, these associations are plotted as mean scores for explanation category, separately for sheep, goats, and intermediates.

GENERAL DISCUSSION

This study shows that the well-known general distortions in people's representation of chance, such as lack of patterns and balancing of alternatives, are indeed positively correlated with paranormal belief. Yet, in all cases but one (avoidance of repetitions in the dice-rolls problem) this association disappears in educated individuals. The correlation between coincidence frequency in everyday life and paranormal belief is, on the contrary, *independent* of the educational level. Moreover, there is no correlation between distortion in the representation of chance and subjective frequency of coincidences.

These results argue against the idea (see the 'chance baseline shift' of Blackmore and Troscianko, 1985, and similar suggestions by Brugger *et al.*, 1990, 1994) that people interpret unusual coincidences as paranormal *because* they underestimate the probability of their occurring by chance. Models based on chance-representation biases contend that, if such probability is underestimated, coincidences will, as all non-chance events, require (and receive) an explanation, a paranormal one if normal explanations do not apply. They will be given a meaning; they will be noticed and remembered. But if coincidences are indeed experienced more often by those who have a more distorted notion of chance, a positive correlation at least as robust as that between coincidence frequency and paranormal belief should obtain between coincidence frequency and estimators of chance misrepresentation, such as avoidance of repetitions in random sequences or errors in probabilistic reasoning. The data reported in this paper carry no indication that such association exists.

The results of the present research are instead consistent with an opposite mechanism. Compared to disbelievers, believers in the paranormal are more prone to seeing meaningful patterns in visual noise (Brugger *et al.*, 1993, Blackmore and Moore, 1994), and tend to need less objective evidence in order to experience cause–effect relationships between temporally contiguous events (Brugger and Graves, 1997). Brugger (1997) has condensed these data in the notion that believers in the paranormal may have a lower 'threshold of causal attribution'. Individuals in which the inclination to connect events causally is present in higher degrees will experience a larger number of meaningful coincidences as a direct consequence; there is no need to assume any prior estimation of coincidences' 'real' probability of occurring by chance. This is in harmony with the finding that sheep and goats are equally good at estimating the chances that a certain coincidence will be found in the general population (Blackmore, 1997). It also explains why, whereas the correlation between paranormal belief and estimators of such 'associative readiness' appears to be a consistent and replicable effect throughout the literature, the correlation between paranormal belief and misinterpretation of chance (under the forms of errors in probabilistic reasoning and repetition avoidance in random sequences) is not.

As adjectives, 'chance' and 'random' denote something lacking a definite plan or purpose, happening without observable cause. The range of events that will subjectively be regarded as random necessarily depends on whether or not one is able to see such plan or purpose, to observe the observable cause. By definition, then, weak and strong pattern-finders will have different representations of randomness. (This is consistent with the idea that judging the randomness of a sequence is based on an attempt to mentally encode it: see Falk and Konold, 1997.) The extent to which this difference will be expressed in differential performance in probabilistic problems, however, will be contingent upon the interaction of a variety of modifiers, such as the type and relative difficulty of the problem, on the one hand; the general intelligence, the educational level, the logical and statistical competence of the respondent, on the other.

In a group where the restraining effects of statistical sophistication tend to be lower, such as in the general, non-student population, we can expect 'truly' random events to be, for strong pattern-finders, much rarer: only events maximally representative of the (psychologically) defining features of chance will be regarded as such. All the differences in the probabilistic performance of sheep and goats can be explained in this way. More specifically, oversensitivity to the presence of patterns will lead to more pronounced pattern avoidance in random (that is, purposeless, patternless, meaningless) sequences; and will concomitantly lead to a diminished appreciation of the relation between expected chance distribution and sample size. For 43% of sheep (versus 4% of goats), for instance, obtaining 2 heads in 4 tosses is a better indicator of a fair coin than obtaining 200 heads in 400 tosses. In this case, the pattern to which sheep may be oversensitive is the impeccable symmetry of such a large distribution: a purposeless, planless, meaningless device will not, left to itself, generate perfection. At the same time, a result of 30 heads in 40 tosses contains, for sheep (82%) more than for goats (58%), a pattern of 'head' superiority obvious enough to reject the notion that the coin might be fair (that is, so obvious that the confidence in taking this decision is not further increased when the heads become 300 in 400 tosses).

What is, then, the connection between random sequences, everyday coincidences, and belief in the paranormal? The propensity to connect events, find patterns, search for causes is of course a useful mechanism for survival. The expression of this basic form of meaning seeking will naturally be liable to considerable inter-individual variation. Those who have less of it will be more able to inhibit its indiscriminate application to chance coincidences; they will refrain from attributing a meaning to co-occurrences of events that do not have a natural explanation (or that are not substantiated by further testing, as suggested by Brugger and Graves, 1997). As a consequence, they will less often come across meaningful coincidences and less often have experiences of a paranormal nature. In a strict sense, paranormal experiences are simply a subset of all possible coincidences – a subset whose ill-defined boundaries may cause confusion and dismay: 'We are moving in a borderland shrouded in fog which blurs the frontier between chance and design, between coincidences which appear to us meaningful in a numinous way, and others which are merely an insult to the laws of probability' (Koestler *et al.*, 1973, p. 213). Coincidences interpretable as precognition, telepathy, clairvoyance experiences fall in the holy class; runs and clusters, small-world encounters, part of the spontaneous associations in the class of insults. If it were paranormal belief to create 'paranormal' coincidences by virtue of interpretation, one would expect to find differences between stronger and weaker believers in the realm of coincidences of the first class only. The fact that this is not the case seems to strengthen the notion that it is not pre-existing belief that creates coincidences, but coincidences that create and foster belief. Naturally, belief itself will favour, in turn, the labelling of many coincidences as 'paranormal'; but this appears to be a subsidiary point. Coincidences create belief because they expose connections unexplained by reason and science, hint at hidden designs (the 'destiny' explanation), and, above all, provide direct paranormal experiences. It is only too easy to imagine how powerful these experiences can be (I used to be a sceptic, until . . .), when even 'meaningless' coincidences, like the recurring of a newly learned word or encountering someone in improbable circumstances, amount so often to a source of perturbation and wonder.

In this light, extreme scepticism appears biologically unnatural, and only possible in a culture where observed connections must carry the burden of proof before becoming established facts. The condition we are genetically equipped for is the search for bonds, patterns, meanings: paranormal belief represents its byproduct at least as much as the ability to classify, the love for music, or the gift for scientific discovery.

ACKNOWLEDGEMENTS

I am very much indebted to Diana Maschietto, without whose substantial collaboration in data collection this work would have not been possible.

REFERENCES

Bar-Hillel M. 1982. Studies of representativeness. In *Judgment under Uncertainty: Heuristics and Biases*, Kahneman D, Slovic P, Tversky A (eds). Cambridge University Press: Cambridge; 69–83.

Bar-Hillel M, Wagenaar WA. 1993. The perception of randomness. In *A Handbook for Data analysis in the Behavioral Sciences*, Keren G, Lewis C (eds). Erlbaum: Hillsdale, NJ; 369–393.

Blackmore SJ. 1997. Probability misjudgement and belief in the paranormal: A newspaper survey. *British Journal of Psychology* **88**: 683–689.

Blackmore S, Galaud K, Walker C. 1994. Psychic experiences as illusions of causality. In *Research in Parapsychology 1991*, Cook EW, Delanoy DL (eds). Scarecrow Press: Metuchen, NJ; 89–93.

Blackmore S, Moore R. 1994. Seeing things: visual recognition and belief in the paranormal. *European Journal of Parapsychology* **10**: 91–103.

Blackmore S, Troscianko T. 1985. Belief in the paranormal: Probability judgements, illusory control, and the 'chance baseline shift'. *British Journal of Psychology* **76**: 459–468.

Brugger P. 1997. Variables that influence the generation of random sequences: An update. *Perceptual and Motor Skills* **84**: 627–661.

Brugger P, Graves RE. 1997. Testing vs believing hypotheses: Magical ideation in the judgement of contingencies. *Cognitive Neuropsychiatry* **2**: 251–272.

Brugger P, Landis T, Regard M. 1990. A 'sheep–goat effect' in repetition avoidance: Extra-sensory perception as an effect of subjective probability? *British Journal of Psychology* **81**: 455–468.

Brugger P, Regard M, Landis T, Cook N, Krebs D, 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, Graves RE. 1995. The roots of meaningful coincidence. *Lancet* **345**: 1306–1307.

Brugger P, Regard M, Landis T, Krebs D, Niederberger J. 1994. Coincidences: Who can say how 'meaningful' they are? In *Research in Parapsychology 1991*, Cook EW, Delanoy DL (eds). Scarecrow Press: Metuchen, NJ; 94–98.

Clarke D. 1995. Experience and other reasons given for belief and disbelief in paranormal and religious phenomena. *Journal of the Society for Psychical Research* **60**: 371–384.

Cohen J. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd edn). Erlbaum: Hillsdale, NJ.

Falk R, Konold C. 1997. Making sense of randomness: implicit encoding as a basis for judgment. *Psychological Review* **104**: 301–318.

Gallup GH, Newport F. 1991. Belief in paranormal phenomena among adult Americans. *Skeptical Inquirer* **15**: 137–146.

Henry J. 1993. Coincidence experience survey. *Journal of the Society for Psychical Research* **59**: 97–108.

Irwin HJ. 1991. The reasoning skills of paranormal believers. *Journal of Parapsychology* **55**: 281–300.

Kahneman D, Tversky A. 1972. Subjective probability: A judgment of representativeness. *Cognitive Psychology* **3**: 430–454.

Koestler A, Hardy A, Harvie R. 1973. *The Challenge of Chance. Experiments and Speculations*. Hutchinson: London.

Roberts MJ, Seager PB. 1999. Predicting belief in paranormal phenomena: A comparison of conditional and probabilistic reasoning. *Applied Cognitive Psychology* **13**: 443–450.

Schmeidler GR. 1943. Predicting good and bad scores in a clarvoyance experiment: a preliminary report. *Journal of the American Society for Psychical Research* **37**: 103–110.

Thalbourne MA, Nofi O. 1997. Belief in the paranormal, superstitiousness and intellectual ability. *Journal of the Society for Psychical Research* **61**: 365–371.
Wierzbicki M. 1984. Reasoning errors and belief in the paranormal. *Journal of Social Psychology* **125**: 489–494.

APPENDIX A: QUESTIONNAIRE ON COINCIDENCES

1. How often have you, in general, come across curious or meaningful coincidences? Please indicate it on the following scale, where 1 = never, 2 = once or twice, 3 = a few times, 4 = many times, and 5 = very often. *[1, 2, 3, 4, 5]*
2. There are many types of coincidences. How often have you experienced coincidences falling in each of the following categories?

   (a) Series or clusters of names, numbers, or events of the same kind (like coming repeatedly across a word, never heard before, in the space of a few hours) *[1, 2, 3, 4, 5]*
   (b) Spontaneous associations (like thinking of someone and running unexpectedly into that person soon afterwards) *[1, 2, 3, 4, 5]*
   (c) 'Small-world' experiences (like encountering a person that one had not seen in a long time in some very improbable place) *[1, 2, 3, 4, 5]*
   (d) Perception of something distant in space (like worrying about a person at the exact time in which that person is having an accident) *[1, 2, 3, 4, 5]*
   (e) Perception of something distant in time (like having a dream that then comes true) *[1, 2, 3, 4, 5]*
   (f) Unexpected solution of a problem (like meeting a friend who wants to sell his computer exactly when we were looking for one) *[1, 2, 3, 4, 5]*
   (g) 'Guardian-angel' experiences (like not arriving in time at a job interview and then discovering that it has been for the best, because a much better chance, which we would otherwise have missed, turns up) *[1, 2, 3, 4, 5]*

3. Do you think coincidences are due to:

   (a) Pure chance *[Yes, No, Don't Know]*
   (b) Destiny *[Yes, No, Don't Know]*
   (c) Divine intervention *[Yes, No, Don't Know]*
   (d) Extra-sensorial perception *[Yes, No, Don't Know]*
   (e) Intuition *[Yes, No, Don't Know]*
   (f) A physical principle not yet discovered by science *[Yes, No, Don't Know]*
   (g) The fact that everything is connected to everything else in the universe *[Yes, No, Don't Know]*

APPENDIX B: PROBABILISTIC REASONING QUESTIONNAIRE (STUDY 1)

(a) Sampling problems

The first problem was a reformulation of one used in Blackmore and Troscianko's study (1985), and consisted of four questions: three were about sampling, with and without replacement, and the fourth was the classical birthday problem. (How many people would you need to gather in a room to have more than a 50% chance that at least two of them will have the same birthday? Possible answers were 23, 61, and 183.) The score for each subject was the total number of correct answers.

The second problem was Kahneman and Tversky's (1972) classical maternity ward dilemma. This was used to assess sensitivity to sample size, in particular to the fact that a deviation from the mean is more likely in smaller than in larger samples. ('A certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50% of all babies are boys. The exact percentage of baby boys, however, varies from day to day. Sometimes it may be higher than 50%, sometimes lower. For a period of 1 year, each hospital recorded the days on which more than 60% of the babies born were boys. Which hospital do you think recorded more such days?' Possible responses were: the larger hospital; the smaller hospital; the probability is the same. Subjects scored −1 if they chose the first answer, 1 if they chose the second, zero if they chose the third.)

(b) Generation of randomness

There were three problems in this class. The first was a replication of the random-string generation question used in Blackmore and Troscianko's study (1985): it basically asked subjects to write a random sequence of 20 digits from 1 to 5. The score for each subject was the number of consecutive repetitions of the same digit. In the remaining two problems, subjects had to simulate the rolling of a die, by writing the digits from 1 to 6 randomly, so as 'to make the resulting sequence as indistinguishable as possible from that of an actually rolled dice'. (These instructions were the same as those given to the subjects of Brugger *et al*'s, 1990, except that they had to call out numbers, rather than write them down.) In both cases, subjects were asked to fill in 66 empty cells (linearly arranged), thus producing a random string of length 66. The difference between the two problems was that, in the first, subjects were given a strip of cardboard with which they were asked to cover the digits as they wrote; that is, they were not able to check their previous responses. In the second, there was no cardboard, and subjects were free to see what they had already written. About half of the subjects received the visible sequence condition first, the other half received the hidden sequence condition first. The score for each subject was the number of consecutive repetitions of the same digit.

APPENDIX C:
PROBABILISTIC REASONING QUESTIONNAIRE (STUDY 2)

(a) Representativeness bias as applied to sample size

There were two problems in this class. The first was a simplified version of the coin-tossing problem used in Blackmore and Troscianko's study (1985). Subjects were told that there were two coins: one was fair and the other was biased (they were not told *how* it was biased). Nine examples of coin tossing (like: I toss the coin 4 times and get 'heads' 2 times) were presented: each time, subjects were asked to assess whether it was more likely that the tossed coin was the fair one or the biased one, on a 5-point scale where 1 = certainly it is not the biased coin, 2 = probably it is not the biased coin, 3 = chances are the same that it is the biased coin or not, 4 = probably it is the biased coin, 5 = certainly it is the biased coin. The nine examples were grouped into three sections of three examples each. In each section, the number of tosses was 4, 40, and 400 (in this order); the number of 'heads' was 2, 20, and 200 in the first section, 3, 30, and 300 in the second, 4, 40, and 400 in the third, corresponding respectively to a 50%, 75%, and 100% proportion of 'heads'. For the first section, subjects scored 1 if they decreased their rating as the number of tosses increased, zero if they gave the same rating over the whole section, −1 if they increased their ratings or responded incoherently by first increasing and then decreasing or vice versa. For the second and third sections, subjects scored 1 if they increased their rating as the number of tosses increased, zero if they gave the same rating over the whole section, −1 if they decreased their ratings or responded incoherently.

The second task concerning sensitivity to sample size was a slight rewording of Bar-Hillel's (1982) pollster problem. ('Two pollsters are conducting surveys to estimate the proportion of voters in their respective cities who intend to vote in favour of a certain party. Firm A operates in a city of 1 million voters; firm B operates in a city of 50,000 voters. Both firms are sampling 1 out of every 1000 voters. Whose estimate would you be more confident in accepting?') Possible answers were: one should have more confidence in the estimate of Firm A; one should have more confidence in the estimate of Firm B; one should have equal confidence in the two estimates. Subjects scored 1 if they chose the first answer, zero otherwise.

(b) Representativeness bias as applied to random sequences

The first task in this class consisted of a reformulation (with a multiple-choice response, and 70 families in place of 72) of Kahneman and Tversky's (1972) birth-order problem. ('All families of six children in a city were surveyed. In 70 families the exact order of births of boys and girls was GBGBBG. What is your estimate of the number of families surveyed in which the exact order of births was BGBBBB?') Possible answers were: about 10 or less, about 30, about 50, about 70, about 90, about 110, about 130 or more. Subjects scored 1 if they chose the correct answer (about 70), zero otherwise.

The second task was a variant of the dice-rolls problem used in Brugger *et al.*'s study (1990). Subjects were presented with pairs of sequences of digits from 1 to 6, each representing six consecutive dice rolls, and asked to estimate which of the two sequences was more likely. One of the sequences of each pair contained more consecutive repetitions of the same digit than the other. Possible responses were: sequence A is more likely; sequence B is more likely; sequences A and B are equally likely. The correct answer is the third; the score corresponded to the total number of correct answers.

(c) Generation of randomness

Three problems fell in this class. The first consisted of simulating the tossing of a coin, by writing the letters H (Head) and T (Tails) randomly, so as to make the resulting sequence as indistinguishable as possible from that of an actually tossed coin. Subjects were asked to fill in 66 empty cells (linearly arranged), thus producing a random string of length 66. The other two were identical to the visible and hidden conditions of Study 1, and were meant as independent replications.

