Falk, R., Falk, R. & Ayton, P. (2009). Subjective Patterns of Randomness and Choice: Some Consequences of Collective Responses. Journal of Experimental Psychology: Learning, Memory, and Cognition, 35(1), pp. 203-224. doi: 10.1037/0096-1523.35.1.203

SUBJECTIVE RANDOMNESS

Subjective Patterns of Randomness and Choice
Some Consequences of Collective Responses

Ruma Falk Raphael Falk Peter Ayton
The Hebrew University of Jerusalem, City University,
Jerusalem, Israel London, England

November 2007

Abstract

Any individual's response intended to be random should be as probable as any other. However, 3 experiments show that many people's independent responses depart from the expected chance distribution. Participants responding to instructions of chance and related concepts favor the available options unequally in a similar way. Consequently, in hide-and-seek games, hiders converge on certain locations and are thereby detected beyond chance by seekers who share their preferences. People agree on salient and on nonsalient options, both of which are preferred under different instructions and even in the absence of instructions. Group responses strongly correlate under diverse, even opposing (e.g., competitive and cooperative) directions. Apparently, common default tendencies, combining random and aesthetic choices, are only somewhat modified under specific instructions. Maximal agreement with others is obtained by implementing one's own aesthetic preferences. These results broadly replicate in one- and two-dimensional tasks. Implications of the findings, their possible roots, and their connection to constructs from, e.g., game theory and subjective-complexity research, are discussed.

Keywords

Randomness; Hide-and-seek; Competition versus cooperation; Aesthetics; Default tendencies.

People's intuitive judgment or perception of randomness and their performance when asked to generate random sets of stimuli have been investigated in many psychological studies, starting in the 1950s. For a comprehensive review, see Nickerson (2002). These studies employed mainly sequences, most comprising binary symbols, and some comprising ternary or multiple symbols.

A consistent tendency to equate randomness with equal total frequencies of the two symbol types has been documented. As for transition probabilities, a majority of participants of different ages – examined under considerable methodological variations – identified randomness with an excess of alternations between successive symbols relative to chance (e.g., Budescu, 1987; Falk & Konold, 1997; Lopes & Oden, 1987; Wagenaar, 1970a, 1970b). Falk (1975) found comparable effects in perception and generation of both one-dimensional (1D) binary sequences and two-dimensional (2D) grids (see Falk & Konold, Figure 2 and Table 5 for a 2D example and 1D results).

It is important to distinguish between analysis of performance of the individual participant and that of the group. As a group, participants who had been instructed to generate a random sequence of hypothetical tosses of a fair coin, produced sequences in which the proportion of heads stayed far closer to .50 than the laws of chance would predict. Kareev (1992) rightly maintained that the overrepresentation of sequences with exactly the right proportion of heads does not necessarily reflect a flaw in participants' understanding of randomness. This cannot be attributed to a biased conception of the individual. Whoever is asked to simulate a sequence of coin tosses, naturally tries to implement the major characteristics of that random process, one of which is the equiprobability of the two outcomes. Hence, one generates a typical random sequence by producing (closely) equal proportions of the two symbol types. Indeed, half heads half tails are more likely to be obtained by chance than any other division. Put differently, an individual participant is not "to blame" for generating a sequence with proportions of heads close to the expected value. Likewise, a group of participants is not to blame for not reproducing the variability of the expected distribution of all random sequences. The participants are not committed, as a group, to reproduce the sampling distribution of the proportions of heads. Their deviation from the latter is a result of matching the proportion in each individual production with the expected proportion. Yet, the discrepancies between aggregate group responses and random distributions might have significant consequences.

Judging by the relative frequencies of the two symbol types, individual productions conformed closely to chance expectation. But judging by the generated <u>alternation rates</u>, the modal individual response as well as group productions deviated from those expected in a random sequence. Whereas the expected probability of alternation in a random binary sequence and grid is .5 (or negligibly greater in the case of constrained equal totals), people's average subjective preference was a probability of .6 or .7 in both generation and judgment tasks (Falk & Konold, 1997). Bar-Hillel and Wagenaar (1991) maintained that, “though the kind of ‘unpredictability’ that people regard as the epitomy [sic] of randomness differs <u>systematically</u> from the real thing, it does not differ <u>radically</u> from it” (p. 448). Indeed, the frequently observed shift from .5 to around .6 or .7 in generating and perceiving randomness does not seem dramatic. But many nonradical deviations between people’s subjective conceptions of randomness and the “real thing" might have practical consequences. Since the tendency is pervasive and consistent, one should not dismiss the potential implications of the bias in mass behavior as, for example, when many people guess lottery numbers or answers in multiple-choice tests. We examine consequences of many people’s choices that rely on their image of randomness in different conditions. Studies that have systematically explored this issue are rather scarce.

Perceiving overalternating sequences as embodying randomness goes hand in hand with seeing runs of identical outcomes that occur by chance as violating randomness. Thus, random streaks of hits in basketball were taken as evidence for a player's "hot hand" (Gilovich, Vallone, & Tversky, 1985), just as random streaks of wins in casino gambling were interpreted as the working of a wave of "luck" (Wagennar & Keren, 1988). In both cases, the researchers showed that chance successions of identical outcomes had invoked a faulty belief in an increased conditional probability of success following a success. This may incur undesirable behavioral consequences. Ayton and Fischer (2004) listed some other situations of illusory perception of positive autocorrelations between successive outcomes.

Some studies asked for choice of a random location in a planar geometric figure. Lisanby and Lockhead (1991) and Psotka (1978) instructed participants to place a dot inside a square, a rectangle, or some other polygon either “at random” or “in the first place that comes to mind” (Lisanby & Lockhead found that similar results were produced in both cases). Obviously, when choosing randomly, any location is as probable as any other, and no single choice can be considered at fault. However, when superimposing all the individual dots within the given outline, the researchers found that the cumulated responses appeared structured. There were more dots in the upper than in the lower half of a square and the dots tended to cluster near the center and along imaginary diagonals. All in all, “collectively, people do not have a perception of randomness as being a uniform distribution” (Lisanby & Lockhead, p. 101). Other studies required generation of random patterns by choosing different numbers of locations within squares of varying sizes (e.g., Dudley, 1982; Falk, 1975). They too found preference of internal locations at the expense of the vicinity of the perimeter. The same tendency was evident when participants had to guess numbers in a given range (Teigen, 1983). Lacey (1962) and Nickerson (2005) asked participants to draw random chords within a circle. They found that the chosen chords were, on average, longer than those which several legitimate random processes would produce. Shorter chords are closer to the periphery, and their relative avoidance repeats the finding of favoring internal zones.

Kahneman and Tversky (1972) accounted for people's bias in identifying randomness with overalternations, or with too short runs, by their expecting short subsequences to locally represent the global population properties. Kareev (1992) and Rapoport and Budescu (1997) argued that these failures stem from attempts to produce a typical sequence within the constraints of people's short-term memory (see also Kubovy & Gilden, 1991). Falk (1975) analyzed randomly chosen subsequences and subgrids within sequences and grids that had been produced under instructions of randomness. She confirmed their being locally representative. The subsets were marked by overly equal divisions between the two symbols. This was tested and confirmed for sequences by Budescu (1987) as well. In locally representative 2D sets the symbols are <u>homogeneously</u> spread over the area more than in random arrangements. Indeed, an attempt to create an impression of homogeneity was reported by Gemelli and Alberoni (1961), whose participants had to mark the position of 20 drops of rain that fall by chance on a rectangle. Ross and Weiner (1963) found that participants who had attempted to construct a random pattern tended to mark cells distributed with equal density on a grid. When the grid was bisected by either a vertical or a horizontal central axis, the numbers of marks were too equally divided between the two sides. Similarly, Nickerson (2005) concluded that perceptual balance, or moderate uniformity of the density over a given area, is a determinant of the degree to which a 2D pattern is perceived as random. The two somewhat contradictory inclinations, to homogeneity on one hand, and to avoidance of edges on the other hand, seem to live side by side in many participants' productions of randomness.

These tendencies were displayed by the performance of 197 participants of Falk (1975). They were asked to imagine a blind process of shuffling and mark 10 cells in a table of 10×10. The group's productions were unduly symmetrical with respect to the two central axes. They tended to eschew marginal cells in preference of internal ones, and they avoided marking neighboring cells. Hence, as in the case of generating sequences, they ended up overalternating in comparison with random productions. An example of two typical patterns – one human-generated and the other generated by random numbers – is presented in Figure 1. These specimens were selected because of simultaneously embodying the modal values of several random variables (e.g., number of neighboring cells or of cells touching the margins).

<u>Insert Figure 1 about here</u>

Implications of situations in which every individual pattern cannot be criticized for being nonrandom, but the collective distribution departs from random expectation, are of special interest. We touch on two examples.

In the Zenith radio experiment (Goodfellow, 1938) the audience was encouraged to send in guesses of five binary symbols obtained by a random machine in the studio and transmitted by a panel of telepathic senders. Analysis of over a million responses revealed that the number of hits was far in excess of chance for some of the machine-generated targets and far below chance for others, depending on the apparent randomness of the target string.

People often think that because a lottery machine draws at random sequences of, say, six numbers from 1 to 49, they must always look random. In the British National Lottery of January 14, 1995 – 133 players hit the winning combination 7, 17, 23, 32, 38, 42 and they had to divide the prize, whereas four weeks later, a single player hit the winning numbers 1, 7, 37, 38, 42, 46 and collected the whole jackpot (Henze & Riedwyl, 1998, p. 24). Note that the former string was rather homogeneously spread across the range 1-49. It did not get close to the edges, nor did it include successive numbers. Contrarily, the latter string started with 1, got close to 49, and had a pair of consecutive numbers.

We endeavor a controlled study of such situations in this research. Many participants are presented with choice tasks in which they get a chance to implement their ideas of randomness and a few related notions, and we explore possible interactive consequences of their aggregate choices.

Overview of the Research

Rationale

Because of people's shared subjective-randomness image, human-generated patterns meant to be random may resemble each other more than truly random patterns do. Hence, when groups of people attempt producing randomness – whether instructed to do so or figuring that this is the best solution for the problem in hand – the similar deviations of the pooled productions from the expected random distribution might have consequential effects. To test this empirically, we devised several mass experiments in which the tasks called for random behavior: When one has to hide objects in a given space so as to be maximally unpredictable, it is normatively advantageous to determine the hiding locations at random. This strategy involves maximal secrecy by concealing knowledge from a rational opponent who otherwise might exploit it (Nickerson, 2002). Randomization also guards the hiders against their own human frailty, that is, against their notorious tendency to generate patterns (Budescu & Rapoport, 1994; Rapoport & Budescu, 1992). The same applies to determining the positions of the right answers upon preparing a multiple-choice test.

Inspired by Falk's (1975) task of choosing 10 random cells in a 10×10 grid and by Rubinstein, Tversky, and Heller's (1997) "treasure game," we chose a task similar to the game of "battleships," with which most of our participants had been familiar since childhood. One group was assigned the task of hiders and one that of seekers (or shooters). To simplify matters, one had to hide/seek only three "ships" (occupying one cell each) in cells of a grid of 5×5 with no restrictions concerning their locations. The instructions presented a competitive task without mentioning randomness. Subjective-randomness research suggests that if both hiders and seekers choose what they consider being random locations, this would fail the hiders and yield an above-chance rate of detections by the seekers. Hence, we predict that hiders' and seekers' choices would be positively correlated. On the other hand, findings that people's selections in a competitive two-person zero-sum game were more nearly random than when they were explicitly instructed to produce randomness (Budescu & Rapoport, 1994; Rapoport & Budescu, 1992) suggest that the two kinds of distributions would be closely uniform and the hit rate would not markedly surpass chance level.

Rubinstein et al. (1997) investigated the behavior of players in a two-person game. One group hid a treasure in one of four locations, and the other sought it. They found that both hiders and seekers tended to avoid the endpoints (as in Dudley, 1982; Falk, 1975; Lisanby & Lockhead, 1991). This bias produced a positive correlation between hiders' and seekers' choices and gave the seekers a considerable advantage. In a more complex design of hide-and-seek, Kareev and Rania (2001) similarly found that hiders overused center locations and underused side locations, and seekers took advantage of these preferences. It is hard to know whether those players meant to act randomly but fell prey to the common biases or perhaps they deliberately tried to select nonsalient locations.

In an attempt to distinguish between these possibilities, we ran a control noncompetitive condition in which participants were instructed to mark three cells in a 5×5 grid at random. Another condition was cooperative – the task was the opposite of competition: Participants in two groups were asked to try to match their choices of three cells in a 5×5 grid with those of partners in the other group without communication. People can often concert their intentions with others, provided there is some clue for coordination (see Schelling's, 1960, section on Tacit Coordination, pp. 54-58). We therefore expect that the distribution of responses to the cooperative instructions would turn out diametrically different from that of the competitive case.

Rubinstein et al. (1997) also ran a cooperative task of selecting one item out of four. Their participants achieved a substantial level of coordination by selecting a focal item that was distinct in some way from the three others. Bacharach (1993) presented game-theoretic arguments for the focal-point principle: "In coordination problems, choose salient options" (p. 256). A salient option is one that stands out from the rest by its uniqueness or distinctiveness, namely, it is the "odd-man-out" item. In a grid of 5×5, the central cell is evidently unique, and the four corner cells are also distinct from the other 21. It stands to reason to expect these cells to be favored by cooperators and avoided by competitors. This may result in sharp differences between the groups' distributions under the two opposing instructions. Moreover, since competitors would apparently be drawn to nonsalient locations, and there are evidently more nonsalient than salient locations in a 5×5 grid, we predict that there would be more agreement between hiders and seekers in the cooperative condition than in the competitive condition.

Branching Out the Tasks

Quite early in running the pilot experiments (that were presented by Ayton & Falk, 1995, and the results of which are included in the present work), it became apparent that participants in the competitive and random conditions performed rather similarly. Surprisingly, however, the distributions obtained under the competitive and cooperative conditions – whose goals were contradictory – were also positively correlated. All this raised the conjecture that response tendencies might be rather impervious to the instructions. We therefore included another control condition with indefinite instructions. The participants of this group got the same task (marking three out of 5×5 cells) with no further directions.

Several symmetric configurations in the productions under all conditions alluded to possible interplay with aesthetic values. Teigen (1984) reported that students judged patterns that had been created by their peers under instructions of randomness slightly more attractive than comparable patterns generated by random numbers. The intricate relation of chance with artistic activities has been touched upon also in other psychological studies concerning subjective randomness (Lisanby & Lockhead, 1991; Ross & Weiner, 1963) and discussed in writings on art (Richter, 1964/1965; Soby, 1958). Deliberate use of accidental elements in works of art has been made by the Dadaists (B. S. Myers, 1969). The author and artist August Strindberg (1849–1912) was one of the early writers on chance in artistic production (Strindberg, 1894/2001). In an attempt to shed some light on that connection, we gave the same task to another group under <u>aesthetic</u> instructions to produce a pleasant-looking pattern.

The overall frequency of choice of each of the 25 cells certainly reflects the popularity of the cell. However, the interrelations among the three marks confound these choices. To get a net measure of the attractiveness of single cells, we conducted another experiment in which one was required to mark only <u>one cell</u> in a 5×5 grid. This was repeated under the same set of different instructions as the task of selecting three cells.

To find out how far people display similar tendencies across different setups, we included – in addition to the tasks pertaining to 2D grids – an analogous set of 1D tasks of <u>generating a string of binary symbols</u>. The first of five symbols was given and the participants were asked to complete the remaining four. This was repeated under the same five conditions as in the case of 5×5 grids. Commonsense predicts that, in both 1D and 2D tasks, cooperators would match each other's responses and depart from randomness more than respondents in all the conditions. Either randomizers or competitors are expected to get closest to randomness.

For the competitive condition of generating a 1D string, we chose the framework of a test of five items, each of which has two possible options of answers. "Hiders" were given the role of test makers (who strive to render the correct answers unpredictable by position clues) and were asked to determine whether the correct answer would be first or second, for each of the four last items. "Seekers" were given the role of testees who do not know the answers and they were asked to guess and choose either the first or the second answer for these items. Falk (1975) suggested – as a result of her findings concerning perceived randomness – that in composing and taking multiple-choice tests, if both parties rely on their image of chance in choosing the position of the correct answer, their shared subjective conceptions might result in unjustified above-chance success of the testees. Attali and Bar-Hillel (2003) and Bar-Hillel and Attali (2002) took that lead and explored the role of answer position in multiple-choice tests. They found systematic position effects of avoiding the edges and balancing the answer keys in the behavior of both test makers and test takers (as in Dudley, 1982; Falk, 1975; Kareev & Rania, 2001; Lacey, 1962; Lisanby & Lockhead, 1991; Ross & Weiner, 1963). These effects were large enough to have real psychometric consequences. We hypothesized that in the tasks of determining the position of the right answer for the remaining four items and guessing these answers, not all possible 16 ordered quartets would be chosen with equal frequencies. Furthermore, the deviations from uniformity of both hiders and seekers would resemble each other so that the hit rate would be inflated.

Three Experiments

The three experiments posed generation tasks: Experiment 1 (2D:3/(5×5)) required marking three cells in a 5×5 table. Experiment 2 (2D:1/(5×5)) required marking one cell in a 5×5 table. Experiment 3 (1D) required completing a 5-symbol binary string in which the first symbol was given. Each of these experiments comprised the following five conditions: Competitive, Cooperative, Random, Indefinite, and Aesthetic. The instructions of the respective conditions were analogous across the three experiments. In all these cases, we took care to formulate the instructions in operational, concrete language and to avoid using theoretical terms. Thus, in instructing to generate randomness, we asked for blind performance, or for mimicking coin-tossing outcomes, and instead of using the adjective "nonsalient," the participant was asked to find hiding places.

Participants and Procedure

The participants in all the experiments were undergraduate students of psychology, education, sociology and biology, mainly from the Hebrew University of Jerusalem. About a third of the participants in Experiment 1 were from City University, London. There were no indications of differences in the patterns of the responses in the two languages and the data were pooled for the analysis.

The experiments were run in a class setting during several academic years. The students participated voluntarily. The design was between groups – different participants were assigned to different conditions. About one percent of the forms were disqualified because of marking a wrong number of cells or failing to complete the 5-symbol string. Altogether, we collected 3028 regular forms.

The students were asked to read the instructions carefully and respond as required without consulting their colleagues. The experiment lasted 10 to 15 minutes. We briefed the students at the end about the purpose of the research.

Analyses of Results

The outcomes were summarized in figures and tables presenting frequency distributions, hit rates, and several effect-size measures, such as correlations and coefficients of consensus whose interpretations are unequivocal. In all the histograms presenting choice distributions, the bar above participants' modal choice was emphasized in black. Interpretation of the results was based primarily on studious evaluation of the extent of their stability (having been based on big numbers) and consistency across a number of variations of the basic experiment. We refrained from lending spurious credibility to the results via tests of significance (Cohen, 1994; Falk, 1998a; Gigerenzer, 1993). We deem it important to avoid the risk of inducing the "illusion of attaining improbability" (Falk & Greenbaum, 1995, p. 78) – that is, the faulty belief that a statistically significant result makes H_0 very improbable – and the associated presumption that "significant" outcomes are bound to replicate. We validated our conclusions, instead, by several kinds of replications (Falk, 1998b) within the research and similarities with other sources. Repeated results reduce the threat of chance. The few features of the research that did not replicate remain moot and require further examination.

Experiment 1 (2D:3/(5×5))

Method

The total number of participants was 1676. In all the forms, a grid (table) of 5×5 empty cells was available beneath the verbal instructions. The rows and columns of the grid were not numbered. The instructions in the seven forms of this experiment are presented below.

Competitive Condition

> Hiders. You are participating in a competitive game. You are asked to hide three "ships" (each of only one cell) in the grid below. Someone in another group will try to sink your ships by "shooting" three cells in an identical grid.
> Please place your three ships by marking X in three cells of your choice.
>
> Seekers. You are participating in a competitive game. You are asked to "shoot" three "ships" (each of only one cell) in the grid below in an attempt to sink them. Someone in another group has already hidden three ships in an identical grid.
> Please indicate your three shots by marking X in three cells of your choice.

Altogether, there were 674 respondents: 367 hiders and 307 seekers.

Cooperative Condition

Cooperation is mutual. Yet, we started by distinguishing between hiders and seekers so as to retain some analogy to the competitive condition.

> Hiders. Imagine that you have to place three important messages (each in only one cell) in the grid below to be found by your partner. Someone in another group will try to find your messages by checking three cells in an identical grid. You have no way to communicate with your partner so as to inform him or her of your choices.
> Please place your three messages by marking X in three cells of your choice.
>
> Seekers. Imagine that you have to find three important messages (each in only one cell) placed in the grid below. Someone in another group has already placed three messages, in the hope you will find them, by marking three cells in an identical grid. You have no way to communicate with your partner so as to learn

> what were his or her choices.
> Please indicate the three places you will check by marking X in three cells of your choice.

Overall, 409 participants answered the cooperative forms: 169 hiders and 240 seekers.

Random Condition

To clarify the instruction to act randomly, we utilized the image of blindness or closed eyes:

> You are asked to mark three cells in the grid below at random, that is, as if they were blindly arranged (but without closing your eyes).
> Please mark X in three cells of your choice.

The number of participants was 250.

Indefinite Condition

> Please mark X in three cells in the grid below.

When asked by some of the 121 participants how they were supposed to do it, our answer was "as it first comes to mind."

Aesthetic Condition

> Please mark X in three cells in the grid below so as to get a pleasant-looking pattern.

The number of respondents was 222.

Results and Discussion

The triplets marked in all the conditions are presented in Supplementary Material A. In every condition, we added all participants' choices of each of the 25 cells. Were the participants to always determine their three marks at random, the distribution of choices over the grid would be close to uniform, namely, with about equal frequencies throughout. No correlation between the distributions of any two conditions is expected in that case. The frequency distributions may indicate which locations or zones in a grid are favored (disfavored) and the extent of consensus among respondents within conditions.

For referring to specific locations, the columns are labeled A through E and the rows 1 through 5. Thus, every cell can be identified by its pair of coordinates. The coordinates were not labeled in the experimental forms. *N* designates the number of marks in each grid. The number of participants is *N*/3.

Choice Distributions

Hiders versus seekers. Figure 2 presents the frequency distributions of marks of the hiders and seekers under competitive and cooperative instructions.

Insert Figure 2 about here

As expected, both hiders' and seekers' choices in the cooperative condition peaked primarily over C3 and secondly over corner cells. These preferences are evidently due to favoring focal points and to guessing others' choices. However, the relative convergence on B4 in the two competitive groups is somewhat surprising. Hiders preferred that cell and C3 and D4 next to it, and seekers acted similarly, anticipating hiders' choices. People are expected to agree about salient locations (as they did in the cooperative case), but they agreed also in the competitive situation, apparently about what constitutes nonsalient sites.

Paradoxically, as can be seen in Figure 2 (and Figure 3), there was also nonnegligible agreement about favored cells between the competitive and the cooperative conditions. This resulted (as shown later) in overall positive correlations between competitors and cooperators. A certain subset of locations seems to be favored and to attract people independently of the two (opposing) instructions. This finding is reminiscent (despite differences in the interpretation) of Shafir's (1993) finding that "some options are both better and worse than others."

The competitive experiment was conducted in two stages: the first, run with 318 participants, was considered "original" and the second – applying identical method and run with 356 participants from the same population – was considered "replication." The participants in each stage were divided about equally between hiders and seekers. The purpose of this arbitrary division into stages was to get a rough idea of the stability (or, conversely, the extent of sampling fluctuations) in the results. The correlation coefficient, across the 25 cells, between the original and the replication of all the competitive choices, was .89. The correlation between all the hiders and all the seekers of the two stages was .84. This means that hiders and seekers responded quite similarly, not unlike two independent groups that perform the same task. Therefore, we combined hiders and seekers into one comprehensive group of 674 competitive participants for comparisons with other conditions. Likewise, since the correlation between hiders and seekers in the cooperative condition was .96, they were pooled into one group of 409 cooperative participants.

All major conditions. In Figure 3, the frequency distributions of all competitors and all cooperators are presented side by side with those of the other experimental conditions. No distinction between hiders and seekers is made in this presentation. Supposedly, the favorite cooperative choice, C3, should have been eschewed by competitors, and conversely, B4 – the preference of competitors – should have been avoided by cooperators. Yet, this did not happen; both C3 and B4 were rather favored under the two conditions, though with different rankings.

The frequency distributions of the five main experimental conditions should be compared with the distribution of a sample (whose size, 532, is within the range of the experimental groups' sizes) that has been obtained by chance simulation, using random numbers (see bottom right diagram in Figure 3). The marks in the simulation condition were determined by randomly drawing a pair of coordinates three times to locate three cells without replacement, and repeating the process 532 times with replacement. This was done to provide a base line for the visual appreciation of the extent of departure from uniformity in the experimental conditions, since a finite random sample is not supposed to yield a strictly flat rectangular distribution.

Insert Figure 3 about here

At first glance, the nonuniformity of the distributions of the five experimental conditions stands out. The variability among the frequencies in these conditions is manifestly greater than in the random simulation (this impression is verified quantitatively later). Visual inspection of the distributions reveals certain similarities in all the experimental conditions. Specifically, B4 and C3 were the first two choices in the competitive, random, indefinite, and aesthetic conditions. In the cooperative condition, C3 was obviously the most popular and B4 came out fourth in popularity. In addition, marginal cells – barring the four corners – were markedly disfavored by humans in all the conditions, but not by random numbers.

Aesthetic decisions. Figure 4 presents the six most popular triplets chosen by 107 out of 222 participants under aesthetic instructions. Altogether, the participants created 77 different patterns (out of 2300 possible ones). The symmetry in the six favored patterns is conspicuous: Three are symmetric with respect to the central vertical axis and three with respect to a diagonal. Notably, B4 was included in three of the six popular patterns and C3 in four of them. Only 12.2% (27 of 222) of all participants generated nonsymmetric triplets, whereas 2068 of the 2300 possible triplets (i.e., 89.9%) are nonsymmetric.

Insert Figure 4 about here

Correlations Between Conditions

Table 1 presents correlation coefficients (across the 25 frequencies) between pairs of choice distributions under all the experimental conditions. The high correlations in Table 1 confirm the visual impression of considerable resemblance among the responses to all experimental conditions (Figure 3). Remarkably, even the correlation between the competitive and the cooperative conditions, whose goals were contradictory, was .71. Obviously, the hiders' tasks in the two conditions are opposites; and the seekers have to guess nonsalient locations in one case and focal locations in the other. All these positive intercorrelations signify similar preferences of locations within the grid under all the conditions. People appear to make similar decisions when they try to avoid others, as well as when they try to match others' choices or to act randomly. They behave likewise even in the absence of definite instructions (note the high correlations with Indefinite). The frequencies in the 25 cells aggregate the three choices of the participants. Some resemblances and differences between triplets (or patterns) of choices under different conditions will be considered later.

Insert Table 1 about here

Hit Rates

To find out how far the seekers manage to detect the hiders, we paired each hider with each seeker in the competitive and cooperative conditions, and counted the number of matching cells in the two triplets, that is, the number of hits. This number may vary from 0 (no match) to 3 (perfect match). When all the choices are made randomly, the probability distribution of these four values is the hypergeometric (Feller, 1957, pp. 41-42). We counted hits also for pairs in groups that were not divided into hiders and seekers. Every participant within the random, indefinite, and aesthetic conditions, was paired with every other participant in the same group. Hence, in a group of *n* participants there were $n(n–1)/2$ pairs. The results showed that in all the conditions, the proportion of zero hits was lower and that of one through three hits was higher than expected by chance. Table 2 presents an abridged version of this analysis, reporting only the percentages of perfect matches (three hits) and the mean number of hits for each condition (the complete distributions according to number of hits in all conditions are presented in Supplementary Material B). A perfect match is extremely rare under random selection, because there are 2300 different triplets of cells in a 5×5 grid, and 1/2300 = .000435. Compared with the chance base line, the results in the table show that the hiders who had tried to avoid detection and the participants who had tried to act randomly failed their mission: Their rate of being outguessed was about nine times greater than expected by chance. This rate was obviously higher for cooperators, and it was hundredfold greater in the aesthetic condition.

Insert Table 2 about here

Consensus Within Conditions

The less uniform a distribution, that is, the greater the variability among its frequencies, the greater the consensus among the participants in that condition. A uniform distribution represents the case of minimal consensus. In contrast, maximal consensus occurs when all the frequencies concentrate (in equal numbers) in three cells and the frequencies in the remaining cells are zero. The actual variability of the frequencies in a given condition is in between these ends, and it forms a certain percent (between 0 and 100) of the variability in the case of full consensus. This percent measures the extent of consensus in that condition: The coefficient of consensus (*CC*) of a given distribution is computed by dividing the standard deviation of its frequencies by the maximal standard deviation and multiplying the result by 100 (Falk & Lann, 2006). Table 3 presents these coefficients for all the conditions. Evidently, the order of the conditions by their *CC*s and by their hit rates (Table 2) is virtually identical. However, the consensus coefficients supplement the previous analyses from another point of view by giving a direct measure of agreement among participants in each condition. Relative to chance simulation, there is greater consensus within all the experimental conditions, including human attempts to be unpredictable and to act randomly. The agreement among people is highest when they make aesthetic choices.

Insert Table 3 about here

Experiment 2 (2D: 1/(5×5))

Method

This experiment was designed as complementary to Experiment 1 where respondents chose three cells in a 5×5 grid, to see if similar choices are made also when only one cell has to be marked. At the same time, it provides a simpler replication of the previous experiment. The instructions were identical verbatim, but for asking to mark one cell in the 5×5 grid. Participants responded to the same five conditions. A total of 386 participants was divided into 99 competitors – 47 hiders and 52 seekers; 122 cooperators – 53 hiders and 69 seekers; 51 randomizers; 51 in the indefinite condition; and 63 in the aesthetic condition. In addition, a chance simulation, in which 70 cells were chosen by random numbers, was prepared.

Results and Discussion

The analyses follow those of Experiment 1 step by step. Here the number *n* of participants in each condition equals the number of marks in the grid. Each figure and table can be compared with the corresponding exhibit in the case of 3/(5×5). After presenting the results of this experiment, we outline analogies and differences between the findings of the two experiments.

Choice Distributions and Their Analyses

Figure 5 presents the distributions of hiders and seekers in the competitive and the cooperative conditions (cf. Figure 2).

Insert Figure 5 about here

As in experiment 1, the resemblance between the distributions of the cooperative hiders and seekers stands out. However, there is also considerable similarity between the distributions of competitive hiders and seekers; the three most popular choices are the same in both. The correlation between hiders and seekers is .86 in the competitive and .97 in the cooperative condition. We therefore pooled hiders and seekers together into one group of 99 competitors and one group of 122 cooperators for subsequent comparisons.

Figure 6 presents the distributions of the five major conditions along with that of the random simulation (cf. Figure 3).

Insert Figure 6 about here

Visibly, the variability among the frequencies within the five conditions is greater than the variability due to random sampling in the simulation. The choices in the experimental conditions were apparently not made at random. More than that, there is similarity among the experimental distributions in the direction of their deviations from uniformity. For the most part, participants avoided marginal locations and preferred more central cells; as a rule, C3, D4. and B4 were mostly favored.

Table 4 presents the correlations between all pairs of conditions (cf. Table 1).

Insert Table 4 about here

The positive correlations in the table indicate considerable agreement in favoring (disfavoring) certain cells for the various purposes, including attempts to select salient and nonsalient locations.

Insert Table 5 about here

When picking 1 of 25 cells, the probability that two random choices will coincide is .04. The percents of hits (matches) between all pairs of hiders and seekers in the competitive and cooperative conditions, and all pairs of participants within the other conditions, are presented in Table 5 (cf. Table 2). The rate in which competitive seekers outguessed the hiders was 2.5 times greater than chance expectation, and in Random and Indefinite it was fourfold greater. This further attests to people's agreement about what constitutes nonsalient locations. Impressive percents of matches were obtained when attempting coordination and, specifically, when implementing one's aesthetic preferences (above 50%). The coefficients of consensus (*CC*) in Table 6 (cf. Table 3) tally with the visual impression (Figure 6) of convergence of the preferences on certain cells and neglect of others.

Insert Table 6 about here

Compared With 2D:3/(5×5)

By and large, the two experiments produced similar results. Many aspects of the outcomes of the task of 3/(5×5) were replicated in the 1/(5×5) task. Not only cooperative, but also competitive hiders and seekers had similar preferences. The choice distributions of all the experimental conditions differed from uniformity. Participants favored (disfavored) about the same locations in analogous conditions in the two experiments: The single most popular cell in both experiments was B4 in Competitive and C3 in Cooperative, Indefinite, and Aesthetic. Even in Random, where the first choice was B4 in Experiment 1 and D4 in Experiment 2, these two cells are, in fact, of similar standing in the grid – both are noncentral and noncorner and they are symmetric with respect to the central vertical axis.

The correlations between pairs of conditions within each of the two experiments were positive and some of them rather high. And so also were the correlations between the two experiments for the same condition, as presented in Table 7. These correlations indicate a tendency to select the same locations, irrespective of whether one chooses one or three cells.

Insert Table 7 about here

In the two experiments, the conditions are ordered almost identically by the level of their hit rates and by their internal consensus, as measured by *CC*. All these measures are greater than chance expectation (or simulation), and they are minimal for Competitive and maximal for Aesthetic.

One important difference is the greater variability among the frequencies within each condition of Experiment 2 relative to the corresponding condition of Experiment 1 (see Figures 5 & 6 vs. 2 & 3 and consider the difference in the percentage scales on the ordinates). The distributions are much more polarized in the case of marking only one cell. This is corroborated by higher CCs in Experiment 2 (Table 6 vs. 3). As noted, the most popular choices in the 1/(5×5) and 3/(5×5) tasks were the same, but apparently, because in Experiment 1 the participants made three choices without replacement, their second and third preferences diluted the sharp differences between the frequencies of favored and disfavored locations.

The answer to the question whether people's choices in the task of 1/(5×5) resemble those of 3/(5×5) is for the most part in the affirmative. Table 8 lists the favored single cells in the two tasks alongside the favored triplets of the 3/(5×5) task. The agreement about the most popular single cells between the two experiments is impressive. Though the favored triplets were not composed exactly of the corresponding three favored single cells, most of these cells appeared in different combinations in the favored triplets of Experiment 1.

Insert Table 8 about here

Preference of symmetry is indicated in both experiments by the popularity of the central cell C3. This was participants' first or second most prevalent response among the single cells in all conditions. Moreover, the most popular triplets in all the conditions of Experiment 1 were symmetric with respect to either a central axis or a diagonal. In general, symmetric triplets were preferred much above chance, as can be seen in Supplementary Material A (and in Table 13). Symmetry imposes constraints, or dependencies, on the interrelations among the three marks of a triplet. This confounds the frequency distributions over single cells. Thus, the correlations between the 1/(5×5) and

3/(5×5) distributions in the same condition (Table 7), although rather high, were not perfect.

Experiment 3 (1D)

Method

Each of 966 participants responded to one of 12 forms – two alternative forms for each of the following six conditions: competitive hiders, competitive seekers, cooperative, random, indefinite, and aesthetic. The task was to complete a string of five binary symbols, the first of which was given. Respondents had to circle one of the two symbols, A or B, in lines 2 to 5 of a list displayed at the bottom of each form:

1	A	B
2	A	B
3	A	B
4	A	B
5	A	B

A was designated as the first choice and encircled in line 1 in one form of each condition and B in the other form. About half the participants in each condition responded to each alternative form. The instructions for the six conditions are presented below:

Competitive Condition

> Hiders. Imagine that you are teachers who prepare a multiple-choice test for your students. The test comprises five questions, each with two alternative answers, A and B, only one of which is correct. The first question is very easy. All the testees will know the right answer. It is A.
>
> You have to determine whether A or B will be the correct answer for the other four questions.
>
> Please circle the answer you choose as the correct one for each question from 2 to 5.

Seekers. Imagine that you are taking a multiple-choice test comprising five questions, each with two alternative answers, A or B, only one of which is correct. The first question is very easy; you know that the right answer is A. However, you have no idea what are the right answers to the following questions.

You try to guess these four answers.

Please circle the answer of your choice for each question from 2 to 5.

Altogether there were 323 respondents: 162 hiders and 161 seekers.

Cooperative Condition

In this experiment only one group of 161 participants got the task of matching partners' choices.

Imagine that you and a partner participate in a game. Below are five lines. The letters A and B are written in each line. Each of you, separately, is asked to choose one of the letters A or B in each line. Your goal is to make identical choices in as many lines as you can. Before starting the game you were allowed a consultation concerning the first line, and you agreed to choose A. However, you did not coordinate your choices for the next four lines.

Try to determine your choices for the following lines so as to match your partner's choices.

Please circle your choice for each line from 2 to 5.

Random Condition

To clarify the instruction to act randomly, we used the analogy of tossing a coin.

Below are five lines in each of which the letters A and B are written. The letter A is circled on the first line.

Complete marking one of the letters, A or B, in the following lines so as to obtain a random order among the lines, that is, as would have been obtained by tossing a coin (but without tossing a coin).

Please circle your choice for each line from 2 to 5.

The number of participants was 160.

Indefinite condition

> Below are five lines in each of which the letters A and B are written. The letter A is circled on the first line.
>
> Please circle one of the two letters, A or B, in each of the following lines, from 2 to 5.

When asked by some of the 163 participants how they were supposed to do it, our answer was "as it first comes to mind."

Aesthetic condition

> Below are five lines in each of which the letters A and B are written. The letter A is circled on the first line.
>
> Complete the marking of one of the two letters, A or B, in the following lines so as to get a pleasant-looking pattern.
>
> Please circle your choice for each line from 2 to 5.

The number of participants was 159.

Results and Discussion

In the analysis, we disregarded the identity of the first circled symbol (whether A or B), denoted it by 1, and recorded each subsequent choice as 1 when it equaled the first one and as 2 when it differed from it. Hence, two reciprocal strings, such as BABBA and ABAAB, in which A and B swap roles, were both encoded 12112. Previous research (e.g., Goodfellow, 1938, 1940) has shown that people tend to prefer one of two apparently equivalent alternatives – such as heads or tails, white or black, circle or cross – in various choice tasks. By starting half the strings with A and half with B, and combining patterns and their reciprocals, we counterbalanced the possible effect of letter preference, to get the net effect of pattern.

There are $2^5=32$ different combinations of five binary choices. When the first choice is denoted 1, the number of ordered quintuplets reduces to 16. In every condition, we summed participants' choices of each of these 16 strings.

Choice Distributions

Table 9 presents the frequency distributions of the choices of strings in all the conditions, alongside Goodfellow's (1938) distribution (in percentages) aggregating approximately 20,000 responses of people who had sent their impressions of a telepathic message of five randomly determined binary symbols. The distribution of a simulation generated by random numbers is given for comparison. The strings are numbered 1 through 16 and ordered lexicographically as by Goodfellow, to enable juxtaposing the results. Boldfaced numbers are the three highest frequencies in each of the human-generated distributions.

Insert Table 9 about here

At a glance, the distributions of competitive hiders and seekers, though not identical, are similar (cf. Figures 2 & 5). The differences among the frequencies are visibly greater within all the experimental distributions than within the much more homogeneous simulated chance distribution (cf. Figures 3 & 6). In each condition, including Competitive, in which game theory's a priori optimal strategy is to act randomly, and Random, where they had been explicitly instructed to do so, participants converged on certain strings and neglected others. These heterogeneous distributions resulted in greater degree of internal consensus in all the conditions, relative to the chance simulation, as quantified by the coefficients of consensus – see bottom line in Table 9. Among the experimental conditions (barring Goodfellow), *CC* is again minimal in Competitive and maximal in Aesthetic (cf. Tables 3 & 6).

Relations Between Conditions

Goodfellow's (1938) instructions matched those of Random (i.e., coin-tossing procedure). However, his results should also be compared with Competitive, since participants would supposedly attempt randomness in this condition as well. Two strings (nos. 7 & 14), preferred in Random, and all three (nos. 6, 7, & 14), preferred in Competitive, were shared with the favorites of Goodfellow's telepathic recipients (Table 9). These strings split between two and three of the two symbols and they alternated either two or three times out of the four transitions between successive symbols. None of them was symmetric, contrary to Aesthetic where symmetric patterns (nos. 5, 11, & 15) were greatly preferred. This resulted in low correlations between Aesthetic and both Competitive and Goodfellow's, as can be seen in Table 10.

Insert Table 10 about here

Despite the prominence of positive correlations, Table 10 contains, for the first time (cf. Tables 1 & 4), some close-to-zero (including negative) correlations between several conditions. The a priori expectation of a negative correlation between the competitive and cooperative responses was not borne out, but, unlike the case of the two 2D experiments, it was low in this experiment. Likewise, Goodfellow's respondents, who conceivably had attempted randomness, as indicated by their positive correlations with our competitive and random conditions, produced low correlations (in absolute values) with our cooperative and indefinite conditions.

Some differences in the factors affecting the choices in the one- and two-dimensional tasks might account for the different intercorrelations patterns. In 5×5 grids, for reasons that we partly understand, B4 and C3 were relatively preferred in all the conditions, thus contributing to the high correlations (Tables 1 & 4). However, in producing a binary string, considerations such as the relative proportions of the two symbols and the number of alternations between them seemed to play some role. For instance, it makes sense that 11111 was cooperators' first choice (Table 9), however, it defies people's image of a representative random string that should comprise the two types of symbols. Hence, it was extremely avoided by Goodfellow's respondents and relatively by our competitors, thus contributing to low correlations with cooperators in both cases (Table 10). At the same time, participants in all the conditions generally agreed on avoiding some of the strings that were divided into one and four symbols and were nonsymmetric (e.g., nos. 2, 3, & 9). This partly explains the absence of extreme negative correlations.

Hit Rates

Insert Table 11 about here

When people act randomly, the probability of a match between two choices is 1/16 = .0625. We crossed the strings produced by every hider and every seeker in Competitive, and, separately, those of every two respondents within the other conditions. The percentages of hits (matches) that were found are presented in Table 11. The maximal hit rate (35%) was obtained in Aesthetic, but all the hit rates exceeded chance expectation. This means that in guessing one of the two answers to each of a four-item test – prepared according to the tester's hiding decisions – the testee might succeed beyond what is warranted by his/her knowledge. Likewise, had Goodfellow's situation been similar to that of the test makers and takers – that is, were both senders and receivers generating their strings according to their subjective randomness – an above-chance telepathic success would have been found. But the excessive hit rate in Goodfellow's case pertains only to matches between the receivers, not between receivers and senders.

Connections With Memory

In discussing the results of Experiment 3 among ourselves and with colleagues, we found ourselves time and again forgetting the exact composition of the most popular string in Competitive and Random (12212) as well as that of Goodfellow's (11212). At the same time, we could easily reproduce from memory the favorite Cooperative and Aesthetic strings (11111 & 12121, respectively). Our memory failure triggered a supplementary, pilot experiment designed to see whether that experience would replicate, and whether the negative relationship between the memorability of a pattern and its perceived randomness (as found by Falk & Konold, 1997) would be confirmed.

Seventy nine new participants read a brief description of Experiment 3 and were shown only the most favored quintuplets that had been chosen under different conditions. Then they were tested for their memory of these strings. Indeed, the two best recalled strings were 11111 and 12121, and the greatest rate of memory errors was obtained for 11212 and 12212. This suggests that hard-to-remember and subjectively-random may be related.

Integrative Comparisons

Outcomes that replicate across several experimental variations and different contexts usually validate each other and extend the generality of the conclusions. On the other hand, divergent findings call for further experimentation and suspension of conclusions. We look first at the extent of consistency among our three experiments and then at comparisons with several external reports.

Experiments 1–3

The first obviously consistent experimental result is the nonuniformity of the aggregate response distributions, whether in 1D or 2D tasks (Figures 3 & 6 and Table 9). The distributions under Cooperative or Aesthetic were a priori not expected to be random. However, the departure of group responses from uniformity under Competitive, Random, and Indefinite was not self-evident beforehand and it was repeated in the three experiments. Another commonality of the three experiments is the high correlations between hiders and seekers (Tables 1, 4, & 10). As a result, the hiders were outguessed more often than they would have been had they been acting randomly as normatively prescribed (Tables 2, 5, & 11).

The extent of deviations from uniformity varied considerably among the conditions within each experiment, as did the internal agreement among participants. However, there was near-perfect agreement between the experiments in the order of the conditions according to the size of their *CC*, as shown in Table 12. This stable order makes a lot of sense on the whole; the one surprising result is the greater consensus among participants in their personal aesthetic choices than in their deliberate cooperative attempts to match each other's responses.

Insert Table 12 about here

Symmetric choices are of special interest because symmetry characterizes pattern, as opposed to randomness, and it ties in with aesthetics. Table 13 presents the percentages of symmetric patterns chosen in the different conditions of the three experiments. In the 1/(5×5) task, the center of symmetry, C3, was selected above the chance rate of 1/25 (i.e., 4%) under all conditions. Likewise, participants produced symmetric triplets exceeding the random expectation of 10% in all the conditions of Experiment 1. In all three experiments, symmetry peaked first under Aesthetic and second under Cooperative. In some conditions the rates of symmetric choices were rather similar across the three experiments. The competitive and random conditions of Experiment 3 are exceptional in their rate of symmetric strings that dropped below or close to the chance level of 25%. This probably accounts for the low correlation between competitors and cooperators in the 1D case (Table 10). Though the near-chance level of symmetric choices in Random might suggest that respondents acted randomly, this was not the case, as evidenced by the inflated consensus in that condition. The frequency distribution of 1D Random deviated considerably from uniformity, and the four symmetric strings (nos. 1, 5, 11, 15) were chosen with divergent frequencies (Table 9).

Insert Table 13 about here

The pattern of intercorrelations among the conditions was the most important feature of the results that was not replicated across the experiments (Tables 1, 4, vs. 10). String specifications, such as the proportions of the two symbols, which were not relevant in the 2D tasks, might partly explain the difference in the pattern of correlations between the 1D and 2D cases. Another important difference between the two cases is that correlations between conditions in the 2D tasks were based on choices of single cells, whereas correlations between 1D conditions were based on choices of complete strings, namely, on patterns. A more appropriate comparison would have been that of correlations based on 1D strings with correlations based on 2D triplets. The latter are understandably lower than correlations based on single cells. However, because in the 1D experiment, a participant had to choose one of 16 different strings, whereas in the 2D task of 3/(5×5) there were 2300 different triplets to choose from, the correlations between conditions within the two experiments are not directly comparable. Moreover, computing intercorrelations based on triplets is impractical. Despite the large number of participants in the different conditions of Experiment 1, their distributions over 2300 possible options contain many isolated outliers of single triplets interspersed among many triplets with zero frequency. Yet, our data suggest that the correlations between conditions in choosing 2D triplets, though lower than the corresponding correlations in choosing single cells (in Table 1), would still be mostly positive. A scrutiny of the choice distributions of triplets in the various conditions reveals that in virtually every pair of conditions there is a nonnegligible overlap between the subsets of the most preferred triplets (see Supplementary Material A).

At the same time, the preponderance of positive correlations (some rather high) is still notable in the 1D results (Table 10). Further scrutiny of the specific intercorrelations reveals that in all three experiments, Indefinite is the condition whose mean correlation with all the other conditions is the highest. This tentatively suggests that the overall similarity between the performances under different conditions reflects a general response tendency that is somewhat modified by differences in the instructions.

This Research and Other Sources

Preference of 1D Strings

The affinity between the results reported by Goodfellow (1938) and those of our 1D Competitive and Random is evidenced by the similar response distributions and the substantial positive correlations. The string 12212 (no. 14 in Table 9) – first choice under Competitive and Random and third in Goodfellow's condition – fits Kahneman and Tversky's (1972) characterization of a subjectively representative random sample. The proportions of the two symbols are as close as possible to equality, and the three alternations preclude runs of length three and at the same time maintain the apparent irregularity of the random process. Kahneman and Tversky ventured that among all possible sequences (disregarding label) of six tosses of a coin, only HTTHTH appears really random. If one ignores the sixth outcome, one gets our favorite string 12212. Konold, Pollatsek, Well, Lohmeier, and Lipson (1993) presented 86 participants with four different ordered quintuplets of Hs and Ts and asked which of them is the most likely result of five flips of a fair coin. Although a majority answered (correctly) that all the strings are equally likely, THHTH (i.e., 12212) got the highest vote among those who had chosen one string.

It stands to reason that 5-symbol strings that include all four of the ordered pairs 11, 12, 21, 22 embody randomness. Pincus and Singer (1996) suggested quantifying the closeness to randomness of a sequence by a measure of approximate entropy (ApEn) based on such considerations. For sequences of length five, their ApEn is maximal for 11221 and 12211. The results in Table 9 show that 11221 (no. 7) was among the favorites, but not the first choice, in the competitive, random and Goodfellow's conditions, whereas 12211 (no. 13) was relatively preferred only in Goodfellow's case. The small but consistent difference between the so-called "objective" and subjective randomness is reflected by the difference between these two strings and 12212 (no. 14) – the first choice of our participants. Whereas sequences 7 and 13, whose ApEn was maximal, alternated only twice (out of four opportunities), sequence 14 had three alternations, in accord with the stable finding of overalternations in sequences perceived as most random. Although string no. 14 (12212) includes only three of the four ordered pairs, the greater number of alternations apparently suffices for portraying an image of randomness. The first choice of Goodfellow's enormous audience – 11212 (no. 6) – also comprised three different ordered pairs and alternated three times.

The above finding that strings chosen as most random were also hardest to memorize agrees well with Falk and Konold's (1997) thesis that the perceived randomness of a sequence is closely associated with its subjective difficulty of encoding. They found high correlations between the length of a tentative condense description of a sequence, measures of difficulty in reproducing the sequence, and its rated randomness. This connection may shed some light on the underlying mechanisms behind people's conception of randomness. Future research devoted to this issue could have people memorize (reproduce) patterns of cells in 2D grids. We hypothesize that the relation between the difficulty of memorizing a pattern – which is the subjective counterpart of its objective complexity (Kolmogorov, 1965) – and the judged randomness of that pattern will be positive.

2D Subjective Patterns

Some important aspects of the results of the two 2D experiments tally with sources that report performance in spatial tasks of producing randomness. The latter came in diverse designs, including choices from 1D and 2D configurations such as squares, some other polygons, or circles.

In a graduate project on subjective randomness, Yaniv Mor (personal communication) conducted a replication of our competitive, random, indefinite, and cooperative conditions. His 244 participants had to mark two cells in a 4×4 grid (a 2/(4×4) task). All his collective distributions departed considerably from uniformity. B3 was by far the most popular choice. The paired correlations of the productions under all conditions were high. The three intercorrelations between Competitive, Random, and Indefinite were all around .9, and that of Competitive-Cooperative was .65. Indefinite was the condition with maximal mean correlation with other conditions. Participants preferred the 4 internal cells to the 12 peripheral ones.

Preference for the internal part of the given frame at the expense of the margins (barring the corners in some cases) is shared by the responses to all the conditions in Experiments 1 and 2 (Figures 3 & 6) and by many other investigations (Attali & Bar-Hillel, 2003; Dudley, 1982; Kareev & Rania, 2001; Lacey, 1962; Nickerson, 2005; Rubinstein et al., 1997; Teigen, 1983). Henze and Riedwyl (1998, Figure 4.3) present the above mentioned two winning sextets in the British National Lottery of 1995 as they were visually arranged in the lottery ticket in a rectangle of 5×10 (without the bottom-right corner cell). Remarkably, all the cells of the most popular choice, made by 133 players – 7, 17, 23, 32, 38, 42 – were internal, and each player won £122,510. In contrast, the winning combination 1, 7, 37, 38, 42, 46, chosen by a single player, had four internal cells and two marginal ones. This player won £10,350,387 (p. 24).

Insert Table 14 about here.

Table 14 presents the distributions of the marks over different regions of the 5×5 grid obtained under random instructions in Experiments 1 and 2 and in Lisanby and Lockhead (1991, Figure 1). We superimposed a grid of 5×5 on the square presenting the aggregate results of Lisanby and Lockhead's participants, each of whom had to mark one dot "at random." Only 168 discernible dots were counted (though the authors reported running 195 participants). As can be seen, the division of participants' choices between central and peripheral cells departed from random expectation in favor of internal cells in all three experiments. In the periphery, noncorner cells were particularly avoided. The extent of agreement between Lisanby and Lockhead's results and those of our Experiments 1 and 2 is further highlighted by the correlations with our Competitive, Random, and Indefinite outcomes, computed across the 25 cells. These were, respectively, .78, .77, and .81 for the 3/(5×5) task, and .87, .83. and .87 for the 1/(5×5) task.

In 10×10 grids, there are 36 marginal and 64 internal cells. In drawing 10 cells at random out of these 100 cells, the probability of getting only internal cells is $C(64,10)/C(100,10)=.00875$. An extreme rate of collective avoidance of the margins was reported by Falk (1975): 25 out of 197 participants, that is, a proportion of .127, who had to randomly mark 10 cells in a 10×10 grid, chose only internal cells. Their rate of avoiding the margins was 14.5 times greater than had they been acting randomly. Interestingly, Christenfeld (1995) reported that people were attracted to the middle of the span of options and were reluctant to choose the ends also in mundane situations – such as choosing a product from a grocery shelf, or a bathroom stall to use – without being instructed to act randomly.

Another point of commonality between Falk's (1975) results of 10/(10×10) (Figure 1) and the present results of 3/(5×5) is a distinct tendency to avoid marking neighbors. Falk reported that 67% of 197 participants generated patterns without even one shared side between any of the 10 marked cells, in contrast to only 16% such patterns in a random simulation. In our 3/(5×5) task, the percents of triplets with no common sides were 82.5 in Competitive, 89.6 in Random, and 84.3 in Indefinite, as compared with 64.8% chance expectation (out of 2300 possible triplets). This is but another way of saying that the subjectively random productions were marked by overalternations.

People's tendency to avoid neighbors when attempting randomness is also displayed in their lottery choices. Our analysis of the numbers chosen by UK lottery players in the 1161 draws held between November 1994 and February 2007 (http://lottery.merseyworld.com/) revealed that on the 560 occasions when the six winning numbers included consecutive numbers, jackpots were shared by fewer people (mean = 2.17) than on the 601 occasions when no consecutive numbers were drawn (mean = 3.38). As a result, average individual winnings were substantially higher on the former occasions (£4,303,329) than on the latter (£3,511,015).

The Dadaist artist Jean Arp (1887–1966) executed drawings and collages seemingly composed of chance arrangements of blots or patches on a smooth background. We bring Arp up as typical of the school of incorporating randomness in artistic work. The story goes (Richter, 1964/1965, p. 51) that, dissatisfied with a drawing he had been working on for sometime, Arp tore it up. When the pieces fluttered to the floor, he was struck by the patterns they formed and accepted this challenge from chance. Figure 7 presents sketches of two of Arp's chance works created four years apart. The sketches preserve only the geometrical relations of the components. Grids of 10×10 were superimposed on these schemes. Though there are 36% marginal cells in a 10×10 grid, we found that less than 5% of the total area of the patches in each of the two schemes lies in marginal cells. Peripheral locations were thus avoided in Arp's pieces as in our participants' choices under random instructions and in the many above-cited studies. In all the conditions of our 2D Experiments, the choice frequencies in the upper half of the grid (disregarding the middle row) were greater than in the lower half. A similar check for left-right halves showed preference of left for all conditions of Experiment 1. Likewise, both Lisanby and Lockhead (1991) and Psotka (1978) found preference of the upper half. In Arp's two pieces, both the upper and the left halves were preferred. Thus, in Figure 7a the total patch area was .28 of the whole area, whereas in the upper-left quartile of the rectangle this proportion was .36; and in Figure 7b the corresponding proportions were .22 and .31.

<u>Insert Figure 7 about here</u>

All in all, it seems that Arp's chance arrangements are a bit "too random." One might suspect (though it would not stand in a law court) that he "improved" the pattern of the fluttering pieces a little. Indeed, in an encyclopedia of art entry ("Arp", 1971, Vol. 1, p. 80) his chance arrangements are said to "exhibit a firm sense of architectonic control." These works would probably be better entitled "According to the Laws of Subjective Chance."

<u>Hiders Versus Seekers</u>

The highly correlated competitive hiders and seekers were united into one competitive group for comparisons with other conditions. Still, in the 2D experiments, the distributions of seekers' choices appear more polarized than those of the hiders (Figures 2

& 5). This would mean that hiders got closer to randomness than seekers, whose agreement in favoring or disfavoring certain locations was greater. The *CC* measures in Experiment 1 were 11.5 for hiders versus 22.6 for seekers, and in Experiment 2 they were 20.2 versus 36.6, respectively. In addition, the percent of triplets with no shared sides was somewhat lower for the competitive hiders of Experiment 1 (79.6) than for the seekers (86.0). The correlation coefficients between the two (arbitrarily divided) original and replication competitive groups in Experiment 1 were .64 for hiders and .91 for seekers. The greater stability of choices of the seekers suggests that their performance was farther removed from randomness.

A similar difference between hiders and seekers was reported by Kareev and Rania (2001). Likewise, in five of six competitive games reported by Rubinstein et al. (1997) there was higher consensus among seekers than among hiders. We established these differences by computing *CC* for the data in their Table 2. Furthermore, they ran also discoordination games in which players had been rewarded for selecting different items from their opponents, much like under hiding instructions. The results showed that in most cases these choice distributions did not depart significantly from random selection. However, the same difference did not hold in our 1D Experiment 3, where *CC* of the competitive hiders was somewhat greater than that of seekers (Table 9). The question of the difference between competitive hiders and seekers is not entirely settled, as opposed to the general similarity in the performance of the two groups that has been consistently confirmed. Further research concerning reasons for possible differences between hiders and seekers is needed. .

Randomizers Versus Competitors

Previous studies (Budescu & Rapoport, 1994; Rapoport & Budescu, 1992) found that when people participate in competitive two-person zero-sum games with rewards, they succeed better in approaching randomness than when explicitly instructed to produce randomness. Also in Experiments 1–3 and in Mor's 2/(4×4) task, participants' productions matched randomness better under Competitive than under Random. This was confirmed by the measures of consensus in all cases as well as by the rates of triplets with no shared sides in Experiment 1. At the same time, departures from randomness still characterized the performance of competitors. The competitive element in our experiments was more hypothetical than practical, yet even in Budescu and Rapoport's studies that involved material consequences the notorious biases were not completely extinguished. The same was true for Rubinstein et al.'s (1997) competitive games that included rewards and penalties.

General Discussion

Mass Productions and Their Consequences

When many individual human choices – each as probable under random selection as any other possible option – are combined, they form nonuniform distributions, contrary to what is expected of many independent random responses. This occurs when people try either to evade or to match others' choices, when they attempt randomness or aesthetic patterns, and in the absence of any instructions. Moreover, these nonuniform distributions are (surprisingly) rather similar to each other, despite different goals set by the instructions. The departure of the distributions from uniformity is more pronounced under aesthetic and cooperative than under random, indefinite, and competitive instructions. Competitors get somewhat closer to uniformity than randomizers, as established in game-theoretic contexts. Yet, in all cases, the variability among the frequencies within the distribution exceeds that of a random simulation, and these variable distributions bear resemblance to each other.

The deviations of all these distributions from randomness are mostly similar: In accord with numerous judgment-under-uncertainty studies, adjacent and peripheral locations are relatively avoided. People seem to agree on salient locations and on nonsalient ones, and there is an (unexpected) extent of overlap between these two subsets of locations and the favorite choices under random and other instructions. Possibly, it all stems from some default preferences that affect the choices under all conditions. The similarities between the aggregate productions under the different conditions (and hiders and seekers within conditions) are an important result of the research. All paired correlations between conditions in 2D – 3/(5×5), 1/(5×5), and 2/(4×4) – tasks were high, including the ones between competitors and cooperators, where a negative correlation would have been a sensible a priori prediction. The one exception was a low (positive) correlation between these two conditions in the 1D experiment. Note, however, that Rubinstein et al. (1997) reported similarities between competitors' and coordinators' choices in their 1D games.

Consequently, in hide-and-seek competitive tasks, seekers detect the hiders beyond chance. This has not been a priori self-evident (as opposed to the case of cooperation), but it was replicated in the three experiments and in other studies using somewhat different procedures (Kareev & Rania, 2001; Rubinstein et al., 1997). Less anticipated was the thrice-replicated finding that different people's individual aesthetic choices match each other better than their intentional effort to do so.

Strategic Considerations

The implications of the results for strategic decisions are not entirely definitive. Game theory's prescription for two-person zero-sum games is choosing the available options with equal probabilities (Budescu & Rapoport, 1994; Rapoport & Budescu, 1992; Rubinstein et al., 1997). This makes sense, because any pattern in hiding, if sensed by seekers, could be utilized to their advantage. However, potential hiders who learn about seekers' tendencies could improve on randomness by avoiding B4 and hiding in, say, noncorner cells in the margins. On the other hand, seekers who know about hiders' preference of the subjectively nonsalient B4 could maximize their success rate by focusing totally on that cell (in this respect, the seekers' task resembles that of cooperators). Perhaps some participants did go a bit that way, and this would be why in many cases hiders' distributions were less polarized than seekers'. Likewise, in the 1D test situation, knowledge of people's inclinations would induce test makers to eschew the 12212 string of binary answers and test takers to concentrate on this pattern. By the same token, as advocated in many studies on lottery choices:

> One smart thing you can do [is to] pick numbers not likely to be chosen by others who would split the prize with you. Given that any combination of … numbers from

> 1 to 49 is as likely as any other, don’t space your numbers the way most people imagine a random series might look. (Myers, 2002, p. 219)

Examples in which lotto choice patterns compatible with that advice paid off have been mentioned. Evidently, this method works as long as it is employed by few players. Once it becomes public knowledge, it might lose its edge.

It is hard to know how far chains of considerations of the type "I think that you'll think that I think that you …" extend. Hence, strategic decisions may swing back and forth. This seems to bring back home the idea of random choices, which was not embraced by participants in our research and in related experiments.

In mundane situations, people often try to guess other's behavior but fail to go one step further and realize that others might do the same. On the eve of major holidays, many people who wish to avoid traffic jams travel at some unearthly hour or through a God-forsaken route only to find themselves stuck in a long slow line of cars with similar traffic avoiders. This is a real-life counterpart of hiding in B4 and sharing that choice with other hiders and seekers.

One clear-cut strategic lesson from the results concerns the advantage of aesthetic choices in achieving interpersonal agreement. If you want to match without prearrangement somebody's decisions, both of you would better not try hard to guess the other’s thoughts, just choose what you like best yourself. The conjunction of Experiment 3 with the finding concerning memory seems to suggest that if you look for a code of, say, five binary symbols for unlocking your door and you want it to be easily recalled by family members, then 11111 seems to be the answer (though it is unlikely to be adopted by most people). However, if you prefer a code that would be forgotten by whoever happened to glimpse it, you might prefer 12212 or 11212, but then you take a risk that strangers might guess your code.

What Do They Actually Do?

The experiments were primarily designed to learn about the interactive consequences of collective choices. This objective has been met. Secondarily, the results might shed some light on the more theoretical issue of the underlying mechanisms of people's performance. Since correlations do not indicate cause-effect relations, the suggested ideas are bound to be somewhat circular, yet, they are worth considering.

A tentative reading of the majority of positive intercorrelations between conditions (the mean of the 30 intercorrelations in the three experiments was .69) is that people primarily act out their <u>default tendencies</u> and only secondarily adjust their responses according to the specifications of each condition. This view is supported by the finding that in the three experiments (and Mor's project) the mean correlation with all other conditions was maximal for <u>Indefinite</u> – a condition in which participants are left to their own devices. Additional weight is lent to this idea by the positive correlations between the responses to the competitive and cooperative conditions. Since the lack of instructions might reasonably be interpreted as a requirement to act randomly, people's shared image of randomness could be one determiner of the similar choices under all conditions. Indeed, the <u>Random</u>-<u>Indefinite</u> correlation was the greatest among all the pairs of conditions, and <u>Random</u> was the condition whose mean correlation with all the others came next to <u>Indefinite</u>. Apparently, the default tendency incorporates a (non-perfect) conception of randomness that affects responses even under cooperative and aesthetic conditions.

The makeup of this default tendency has yet to be better understood. Next to <u>Indefinite</u> and <u>Random</u>, the condition mostly correlated with all the others (across experiments) was <u>Aesthetic</u>. Tentatively, people's spontaneous responses are affected by their image of randomness as well as by aesthetic preferences. Possibly, "what first comes to mind" sways between "random" and "aesthetically pleasing." There is a lot of affinity between cooperative and aesthetic considerations, but it seems that aesthetic elements creep indirectly into the performance under all conditions, including <u>Competitive</u> and <u>Random</u>, via the general response tendency. This conjecture is supported by the surfeit of symmetric choices in almost all the cases.

There is considerable overlap between the concepts of simplicity and symmetry. The latter is an important constituent of the former (and of aesthetics). According to Chater (1999), simplicity works as a guiding principle in many aspects of cognition. People's default tendencies apparently represent a compromise between their attempt to be random, or unpredictable (hence the prominence of B4), and their penchant for Chater's simplicity principle (hence the surfeit of symmetric patterns).

Why B4? And Why C3?

A major factor in the similarities between all the response distributions in the 2D conditions is the relative prominence of two cells: B4 and C3 (Figures 3 & 6 and Table 8). B4 and D4, its mirror image through the vertical axis of symmetry, are interchangeable and the same arguments pertain to both (we do not consider B2 and D2 equivalent to B4 because of the robust findings showing that people consistently prefer the upper half of a grid). C3 has no mirror image, since it is the center of symmetry. Because C3 stands out, it is the first spontaneous (i.e., indefinite) and aesthetic choice as well as the focal point for coordination. The default popularity of C3 apparently carries over to the remaining conditions (one could object that competitors prefer C3 in order to outsmart their opponents, but this would not explain the cell's popularity in Random). The reasons for the popularity of B4 are, however, less obvious.

The status of B4 in the 5×5 grid resembles that of the number 7 among the 10 digits (Kubovy & Psotka, 1976; Teigen, 1983). Preference of 7 is apparently accounted for by its ostensible lack of distinguishing characteristics due to its position neither at the middle nor at the end of the digit span. More than that, the prime number 7 cannot be factorized (see Griffiths & Tenenbaum, 2001, for attempted mathematical models for predicting choice of numbers and strings of five binary symbols). By the same token, avoidance of the center, corners, and margins (combined with preference of the upper half) leads to favoring B4. This cell seems to epitomize the idea of nonsaliency. The following explanations are somewhat speculative and their feasibility has to be further examined.

Barring the definition of a cell by its pair of coordinates, the shortest verbal description of B4 seems to be longer than that of C3 or corner and margin cells. This is reminiscent of Kolmogorov's (1965) characterization of a sequence's complexity by the length of its shortest computer algorithmic description and of Falk and Konold's (1997) finding of amazingly high correlations between the length of an attempted concise description of a sequence and its judged randomness and other measures of subjective complexity (see Falk & Konold's *DP* index and Table 2). By the same token, the increased difficulty of encoding B4 could presumably account for its greater perceived randomness.

In the following two paragraphs, we mention two alternative approaches to quantifying a cell’s subjective complexity or its perceived randomness. Both have appeared in the literature. These suggested explanations differ from each other. It is hard to know whether either represents people’s experience of randomness or complexity, but they are worth considering.

Suppose one rotates the 5×5 grid stepwise, 90^{o} at a time, and reflects it through the two middle axes and the diagonals. C3 will stay put, whereas B4 will be mapped into D4, D2, and B2. Garner (1970) proposed the number of different patterns obtained under rotations and reflections as a reverse measure of a pattern's perceived "goodness." This number, $N(RR)$, can also be used to quantify the complexity of cells in a grid. $N(RR)$ is 1 for C3 and 4 for B4. Indeed, C3 represents extreme simplicity; it is unique with no alternatives. However, 4 is not the greatest possible $N(RR)$; in 5×5 grids, the maximum is 8 as, for example, for B5, because the rotations and reflections map B5 into another seven cells: D5, E4, E2, D1, B1, A2, A4. Thus, the popularity of B4 in <u>Random</u> and <u>Competitive</u> seems to tally with findings that an intermediate degree of objective complexity is perceived as most subjectively complex or random. The string 12212 – first choice in <u>Competitive</u> and <u>Random</u> in Experiment 1 – is also characterized by less than Pincus and Singer's (1996) maximal measure of closeness to randomness (ApEn). Likewise, the peak of the function of perceived randomness of binary sequences corresponds to sequences with excessive alternations, whose objective complexity is less than maximal (Falk & Konold, 1997).

A different measure that may characterize the complexity of cells is the size of the equivalence set to which a cell belongs. An equivalence set comprises cells of the same $N(RR)$. It can easily be verified that classifying the 5×5 cells according to their $N(RR)$ yields three subsets. The smallest is of size 1 (comprising only C3, whose $N(RR)$=1); the second is of size 8 (comprising cells whose N(RR)=8); and the greatest is of size 16 (comprising cells whose $N(RR)=4$). B4 belongs to the latter. The prevalent choice of B4 in generating randomness agrees with Teigen's (1984) thesis that stimuli judged as random belong to greater equivalence groups than those judged as nonrandom (see Falk & Konold, 1997; and Kubovy & Gilden, 1991). Thus, B4 is considered random as a member of the greatest group of cells of the same $N(RR)$.

B4 is also among the relatively preferred aesthetic options. The relations between perceived randomness of stimuli, their aesthetic appeal, and their (objective or subjective) complexity are complex and fraught with inconsistencies (Lisanby & Lockhead, 1991; Ross & Weiner, 1963; Teigen, 1984). There are, however, some indications that moderately complex stimuli are preferred aesthetically (e.g., Cox & Cox, 2002). This would be consistent with the inclusion of B4 among the aesthetically pleasing choices. Furthermore, the cell B4 contains the point that cuts both the horizontal and vertical sides of the grid into segments that relate to each other by the golden ratio that is known to be aesthetically liked. The same is true for B3, the favorite under all conditions in Mor's 4×4 grid. Gardner (1961, chap 8) maintained that this ratio (that is almost as ubiquitous as the number π) has the pleasant propensity for popping up where least expected. Indeed, B4 appeared among the favorites in most of the response distributions.

Conclusions

People focus naturally on stimuli that stand out as distinctive, and hence they agree about what is prominent. Remarkably, however, a great deal of consensus was also found about what is random or inconspicuous. If in mass situations people try to hide or avoid others according to their intuitive disposition, they run the risk of being overly found out. In lottery terms, if you try to mimic randomness, you'll find that you have too many partners to share the jackpot with.

The similarities between response distributions are not limited to hiders and seekers. They pertain to virtually all the conditions studied in this work. Particularly noteworthy is the extent of agreement between choices of competitors and cooperators. One should bear in mind, however, that these results were obtained in a between-subjects design. Were the same participants responding to both the competitive and cooperative instructions, they would presumably have been sensitive to the reversal of the instructions and responded differently. At the same time, in real situations, one generally encounters only one of these conditions at a time and is rarely required to respond to both, as in a within-subjects design.

Symmetry and aesthetics play an important role in determining people's choices. The aesthetic effect is highlighted in this work by many positive correlations and by the finding that people agree on what counts as aesthetic more than on what is salient, let alone nonsalient. It goes without saying that achieving aesthetic impact is the driving force behind artistic work. Beauty was also described as the ultimate criterion for good mathematical work by the renowned mathematician Hardy (1940). As noted by Chater (1999), scientists frequently report strong aesthetic preferences in theory construction and evaluation. All this suggests that aesthetic values – whether deliberate or not – might weigh heavier in people's cognition and action than heretofore believed.

References

Arp. (1971). In D. Bell (Ed.), *Encyclopaedia of art* (Vol. 1, pp. 80-81). London: Encyclopaedia Britannica.

Attali, Y., & Bar-Hillel, M. (2003). Guess where: The position of correct answers in multiple-choice test items as a psychometric variable. *Journal of Educational Measurement, 40*(2), 109-128.

Ayton, P., & Falk, R. (1995, August). *Subjective randomness in hide-and-seek games.* Paper presented at the 15th bi-annual conference on Subjective Probability, Utility and Decision Making, Jerusalem, Israel.

Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & Cognition*, *32*(8), 1369-1378.

Bacharach, M. (1993). Variable universe games. In K. Binmore, A. Kirman & P. Tani (Eds.), *Frontiers of game theory* (pp. 255-275). Cambridge, MA: MIT Press.

Bar-Hillel, M., & Attali, Y. (2002). Seek whence: Answer sequences and their consequences in key-balanced multiple-choice tests. *The American Statistician, 56*(4), 299-303.

Bar-Hillel, M., & Wagenaar, W. A. (1991). The perception of randomness. *Advances in Applied Mathematics*, *12*, 428-454.

Budescu, D. V. (1987). A Markov model for generation of random binary sequences. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 25-39.

Budescu, D. V., & Rapoport, A. (1994). Subjective randomization in one- and two-person games. *Journal of Behavioral Decision Making, 7*, 261-278.

Chater, N. (1999). The search for simplicity: A fundamental cognitive principle? *The Quarterly Journal of Experimental Psychology, 52A*(2), 273-302.

Christenfeld, N. (1995). Choices from identical options. *Psychological Science, 6*(1), 50-55.

Cohen, J. (1994). The earth is round (*p*<.05). *American Psychologist, 49*(12), 997-1003.

Cox, D., & Cox, A. D. (2002). Beyond first impressions: The effects of repeated exposure on consumer liking of visually complex and simple product designs. *Journal of the Academy of Marketing Science, 30*(2), 119-130.

Dudley, B. (1982). Exploring a random sampling method used in biology. *Teaching Statistics, 4*(2), 53-56.

Falk, R. (1975). *Perception of randomness.* Unpublished doctoral dissertation (in Hebrew with English abstract). The Hebrew University of Jerusalem.

Falk, R. (1998a). In criticism of the null hypothesis statistical test. *American Psychologist, 53*(7), 798-799.

Falk, R. (1998b). Replication -- A step in the right direction. *Theory & Psychology, 8*(3), 313-321.

Falk, R., & Greenbaum, C. W. (1995). Significance tests die hard: The amazing persistence of a probabilistic misconception. *Theory & Psychology, 5*(1), 75-98.

Falk, R., & Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review, 104*(2), 301-318.

Falk, R., & Lann, A. (2006). Consensus is unfair. *Teaching Statistics, 28*(3), 81-83.

Feller, W. (1957). *An introduction to probability theory and its applications* (2nd ed. Vol. 1). New York: Wiley.

Gardner, M. (1961). *More mathematical puzzles and diversions*. Harmondsworth, Middlesex, England: Penguin.

Garner, W. R. (1970). Good patterns have few alternatives. *American Scientist, 58*, 34-42.

Gemelli, A., & Alberoni, F. (1961). Experimental studies of the concept of chance. *The Journal of General Psychology, 65*, 3-24.

Gigerenzer, G. (1993). The superego, the ego, and the id in statistical reasoning. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 311-339). Hillsdale, NJ: Lawrence Erlbaum.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology, 17*, 295-314.

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

Goodfellow, L. D. (1940). The human element in probability. *The Journal of General Psychology, 23,* 201-205.

Griffiths, T. L., & Tenenbaum, J. B. (2001). Randomness and coincidences: Reconciling intuition and probability theory. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 370-375). Hillsdale, NJ: Lawrence Erlbaum.

Hardy, G. H. (1940). *A mathematician's apology*. Cambridge: Cambridge University Press.

Henze, N., & Riedwyl, H. (1998). *How to win more: Strategies for increasing a lottery win*. Natick, MA: A K Peters.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*(3), 430-454.

Kareev, Y. (1992). Not that bad after all: Generation of random sequences. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1189-1194.

Kareev, Y., & Rania, A. (2001, November). *People's belief in their ability to foil competition.* Paper presented at the 42nd Meeting of the Psychonomic Society, Orlando, FL. USA.

Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems in Information Transmission*, *1*(1), 1-7.

Konold, C., Pollatsek, A., Well, A., Lohmeier, J., & Lipson, A. (1993). Inconsistencies in students' reasoning about probability. *Journal for Research in Mathematics Education, 24*(5), 392-414.

Kubovy, M., & Gilden, D. (1991). Apparent randomness is not always the complement of apparent order. In G. R. Lockhead & J. R. Pomerantz (Eds.), *The perception of structure* (pp. 115-127). Washington, DC: American Psychological Association.

Kubovy, M., & Psotka, J. (1976). The predominance of seven and the apparent spontaneity of numerical choices. *Journal of Experimental Psychology: Human Perception and Performance, 2*(2), 291-294.

Lacey, O. L. (1962). The human organism as a random mechanism. *The Journal of General Psychology, 66*, 321-325.

Lisanby, S. H., & Lockhead, G. R. (1991). Subjective randomness, aesthetics, and structure. In G. R. Lockhead & J. R. Pomerantz (Eds.), *The perception of structure* (pp. 97-113). Washington, DC: American Psychological Association.

Lopes, L. L., & Oden, G. C. (1987). Distinguishing between random and nonrandom events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 392-400.

Myers, B. S. (Ed.) (1969). Chance configuration. In *McGraw-Hill dictionary of art* (Vol. 2, pp. 11-12). New York: McGraw-Hill.

Myers, D. G. (2002). *Intuition: Its powers and perils*. New Haven, CT: Yale University Press.

Nickerson, R. S. (2002). The production and perception of randomness. *Psychological Review, 109*(2), 330-357.

Nickerson, R. S. (2005). Bertrand's chord, Buffon's needle, and the concept of randomness. *Thinking & Reasoning, 11*(1), 67-96.

Pincus, S., & Singer, B. H. (1996). Randomness and degrees of irregularity. *Proceedings of the National Academy of Sciences, USA, 93*, 2083-2088.

Psotka, J. (1978). Perceptual processes that may create stick figures and balance. *Journal of Experimental Psychology: Human Perception and Performance, 4*(1), 101-111.

Rapoport, A., & Budescu, D. V. (1992). Generation of random series in two-person strictly competitive games. *Journal of Experimental Psychology: General, 121*, 352-363.

Rapoport, A., & Budescu, D. V. (1997). Randomization in individual choice behavior. *Psychological Review, 104*(3), 603-617.

Read, H. E. (1968). *Arp*. London: Thames and Hudson.

Richter, H. (1965). *Dada: Art and anti-art* (D. Britt, Trans.). London: Thames and Hudson. (Original work published 1964).

Ross, B. M., & Weiner, S. G. (1963). On making a random pattern. *Perceptual and Motor Skills, 17*, 587-600.

Rubinstein, A., Tversky, A., & Heller, D. (1997). Naive strategies in competitive games. In W. Albers, W. Güth, P. Hammerstein, B. Moldovanu & E. van Damme (Eds.), *Understanding strategic interaction: Essays in honor of Reinhard Selten* (pp. 394-402). Berlin: Springer.

Schelling, T. C. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.

Shafir, E. (1993). Choosing versus rejecting: Why some options are both better and worse than others. *Memory & Cognition*, *21*(4), 546-556.

Soby, J. T. (Ed.). (1958). *Arp*. New York: Museum of Modern Art.

Strindberg, A. (2001). On chance in artistic creation. *Cabinet Magazine Online* issue 3. from http://www.cabinetmagazine.org/issues/3/i_strindberg.php (K. Board, Trans. Original work published 1894).

Teigen, K. H. (1983). Studies in subjective probability I: Prediction of random events. *Scandinavian Journal of Psychology, 24*, 13-25.

Teigen, K. H. (1984). Studies in subjective probability V: Chance vs. structure in visual patterns. *Scandinavian Journal of Psychology, 25*, 315-323.

Wagenaar, W. A. (1970a). Appreciation of conditional probabilities in binary sequences. *Acta Psychologica, 34*, 348-356.

Wagenaar, W. A. (1970b). Subjective randomness and the capacity to generate information. *Acta Psychologica*, *33*, 233-242.

Wagennar, W. A., & Keren, G. B. (1988). Chance and luck are not the same. *Journal of Behavioral Decision Making, 1*, 65-75.

Author Note

Ruma Falk, Department of Psychology, and School of Education, Hebrew University of Jerusalem; Raphael Falk, Department of Genetics and the Program for History and Philosophy of Science, Hebrew University of Jerusalem; Peter Ayton, Department of Psychology, City University, London.

The study was partially supported by the Sturman Center for Human Development, Hebrew University of Jerusalem. We thank Yaniv Mor for allowing us to use his data, Yael Oren for helping with the figures, Peter Barr for computer programming, and Richard K. Lloyd for providing UK lottery data. We are grateful to the editor and two reviewers for thoughtful comments and constructive suggestions.

Supplementary material is available online.

Correspondence concerning this article should be addressed to Ruma Falk, 3 Guatemala St., Apt. 718. Jerusalem, 96704 Israel. E-mail: rfalk@cc.huji.ac.il

Table 1

Intercorrelations Between Pairs of Conditions in Experiment 1 – 2D:3/(5×5)

(In parenthesis: number of marks in the grid)

Condition	Random (750)	Competitive (2022)	Cooperative (1227)	Indefinite (363)
Competitive	.85	.84[a]	–	–
Cooperative	.70	.71	.96[a]	–
Indefinite	.91	.84	.83	–
Aesthetic (666)	.88	.76	.82	.88

Note. The expected correlation between two randomly determined distributions is zero.

[a] Between hiders and seekers of that condition.

Table 2

Summary of Hit Statistics Under Different Conditions in Experiment 1 – 2D:3/(5×5)

Condition	No. of pairs[a]	Percent of three hits	Mean no. of hits
Chance expectation	—	.0435	.36
Competitive	112,669	.38	.42
Cooperative	40,560	1.83	.60
Random	31,125	.37	.48
Indefinite	7,260	.95	.56
Aesthetic	24,531	4.35	.70

[a] In the competitive and cooperative conditions every hider was paired with every seeker, and in the other conditions everybody was paired with everybody else.

Table 3

Measures of Consensus (CC) According to Condition in Experiment 1 – 2D:3/(5×5)

Condition	*N*	*CC*
Chance simulation	1596	4.08
Competitive	2022	15.9
Cooperative	1227	30.2
Random	750	21.8
Indefinite	363	28.7
Aesthetic	666	36.3

Note. *N* is the number of marks in the grid, which is three times the number of the participants.

Table 4

Intercorrelations Between Pairs of Conditions in Experiment 2 – 2D:1(5×5)

(In parenthesis: number of participants = number of marks in the grid)

Condition	Random (51)	Competitive (99)	Cooperative (122)	Indefinite (51)
Competitive	.87	.86[a]	–	–
Cooperative	.65	.60	.97[a]	–
Indefinite	.95	.93	.70	–
Aesthetic (63)	.63	.57	.99	.68

Note. The expected correlation between two randomly determined distributions is zero.

[a] Between hiders and seekers of that condition.

Table 5

Hit Rates Under Different Conditions in Experiment 2 – 2D:1/(5×5)

Condition	No. of pairs[a]	Percent of hits
Chance expectation	–	4.0
Competitive	2,444	10.1
Cooperative	3,657	37.9
Random	1,275	16.7
Indefinite	1,275	16.2
Aesthetic	1,953	53.4

[a] In the competitive and cooperative conditions every hider was paired with every seeker, and in the other conditions everybody was paired with everybody else.

Table 6

Measures of Consensus (CC) According to Condition in Experiment 2 – 2D:1/(5×5)

Condition	*n*	*CC*
Chance simulation	70	12.0
Competitive	99	27.9
Cooperative	122	63.9
Random	51	38.6
Indefinite	51	38.0
Aesthetic	63	72.2

Note. *n* is the number of participants = number of marks in the grid.

Table 7

Correlations Between Results of 2D: 3/(5×5) and 1/(5×5) According to Condition in Experiments 1 and 2

Condition	Correlation
Chance simulation[a]	–.14
Competitive	.80
Cooperative	.74
Random	.69
Indefinite	.74
Aesthetic	.75

[a] The expected correlation when all the choices are determined by chance is zero.

Table 8

Most Popular Cells and Most Popular Triplets in 2D: 1/(5×5) and 3/(5×5) Tasks by Condition

Condition	Most popular single cells		Most popular triplets
	Experiment 2 1/(5×5) task	Experiment 1 3/(5×5) task	Experiment 1 3/(5×5) task
Competitive	B4; C3; D4	B4; C3; D4	(B4, C3, D2) (A5, C3, E1) (B3, D2, D4)
Cooperative	C3[a]	C3; A5; E5	(A5, C3, E1) ; (A1, C3, E5)
Random	D4; C3; B4 C4	B4; C3; D4	(B4, C3, D2) ; (B2, B4, D3)
Indefinite	C3; D4; B4	C3; B4; D3	(B4, C3, D2) ; (B2, C3, D4)
Aesthetic	C3[a]	C3; B4; D4	(B3, C4, D3) ; (B4, C3, D4) (B4, C3, D2)

Note. Items are written in decreasing order of their frequencies. Items of equal frequencies are written one under the other.

[a] Other cells were of negligible frequency.

Table 9

Distribution of Choices by String and by Condition in Experiment 3 – 1D

No.	String	Condition: Competitive Hiders	Competitive Seekers	Competitive Total	Cooperative	Random	Indefinite	Aesthetic	*Goodfellow*[a] (%)	Chance simulation
1	11111	4	11	15	***49***	2	16	6	0.84	16
2	11112	4	1	5	1	1	1	0	1.39	11
3	11121	7	4	11	7	4	4	0	3.73	17
4	11122	8	1	9	0	1	0	2	4.48	20
5	11211	9	4	13	3	6	3	***21***	4.99	11
6	11212	***25***	9	***34***	6	13	4	1	***14.23***	14
7	11221	***18***	***20***	***38***	8	***17***	16	5	***11.82***	18
8	11222	5	5	10	5	6	2	0	5.66	19
9	12111	0	3	3	1	2	1	0	3.22	17
10	12112	12	15	27	12	14	***23***	2	8.68	21
11	12121	4	11	15	***37***	***30***	***41***	***88***	4.34	17
12	12122	3	6	9	4	4	8	0	5.70	21
13	12211	12	7	19	6	12	7	5	10.90	19
14	12212	***39***	***39***	***78***	***15***	***40***	***25***	5	***11.66***	19
15	12221	9	9	18	1	5	4	***23***	6.48	18
16	12222	3	***16***	19	6	3	8	1	1.95	18
Total		162	161	323	161	160	163	159	100.07	276
CC[b]		24.6	23.5	22.6	33.9	27.7	27.8	55.3	16.2	4.4

Note. The total number of participants in the experiment was 966.

[a] Based on ~20,000 responses in Table III of Goodfellow (1938). [b] *CC* is a coefficient of consensus, bounded between 0 and 100.

Table 10

Intercorrelations Between Pairs of Conditions in Experiment 3 – 1D

(in parenthesis: number of participants)

Condition	Random (160)	Competitive (323)	Cooperative (161)	Indefinite (163)	Aesthetic (159)
Competitive	.81	.76[a]	–	–	–
Cooperative	.34	.13	–	–	–
Indefinite	.79	.46	.71	–	–
Aesthetic	.46	-.05	.49	.68	–
Goodfellow[b]	.57	.69	-.20	.17	-.10

Note. The expected correlation between two randomly determined distributions is zero. [a] Between competitive hiders and seekers. [b] Based on ~20,000 responses (Goodfellow, 1938, Table III).

Table 11

Hit Rates Under Different Conditions in Experiment 3 – 1D

Condition	No. of pairs[a]	Percent of matching strings
Chance expectation	–	6.25
Competitive	26,082	10.37
Cooperative	12,880	16.51
Random	12,720	12.92
Indefinite	13,203	12.97
Aesthetic	12,561	34.54
Goodfellow[b]	$\sim 19{,}999 \times 10^4$	~8.73

[a] In the competitive condition every hider was paired with every seeker, and in the other conditions everybody was paired with everybody else. [b] Based on Goodfellow (1938, Table III).

Table 12

Rank Order of the Conditions in the Three Experiments According to the Size of Their Coefficients of Consensus

Condition	1D[a] Experiment 3	2D 3/(5×5) Experiment 1	2D 1/(5×5) Experiment 2
Chance simulation	1	1	1
Competitive	2	2	2
Cooperative	5	5	5
Random	3	3	4
Indefinite	4	4	3
Aesthetic	6	6	6

Note. The ranks vary from the lowest (1) to the highest (6).

[a] Goodfellow's ranking is in between chance simulation and the competitive condition.

Table 13

Percents of Participants Who Generated Symmetry According to Condition in the Three Experiments

(in parenthesis: number of participants)

Condition	1D	2D	
	Experiment 3	3/(5×5) Experiment 1	1/(5×5) Experiment 2
Chance level	25	10	4
Competitive	19 (323)	35 (674)	18 (99)
Cooperative	56 (161)	56 (409)	65 (122)
Random	27 (160)	33 (250)	25 (51)
Indefinite	39 (163)	46 (121)	27 (51)
Aesthetic	87 (159)	88 (222)	73 (63)
Goodfellow	17 (~20,000)	–	–

Table 14

Proportion of Choices Under Random Instructions According to Location of the Cells and to Source

(in parenthesis: number of marks in the 5×5 grid)

Source	Location of cells		
	Internal (9 cells)	Noncorner margins (12 cells)	Corners (4 cells)
Chance expectation	.36	.48	.16
Experiment 1 2D: 3/(5×5) (750)	.58	.28	.14
Experiment 2 2D: 1/(5×5) (51)	.92	.04	.04
Lisanby and Lockhead (1991, Figure 1) (168)	.61	.21	.18

Figure Captions

Figure 1. Two typical patterns, meant to be random, obtained by marking 10 cells in a 10×10 grid (from Falk, 1975)

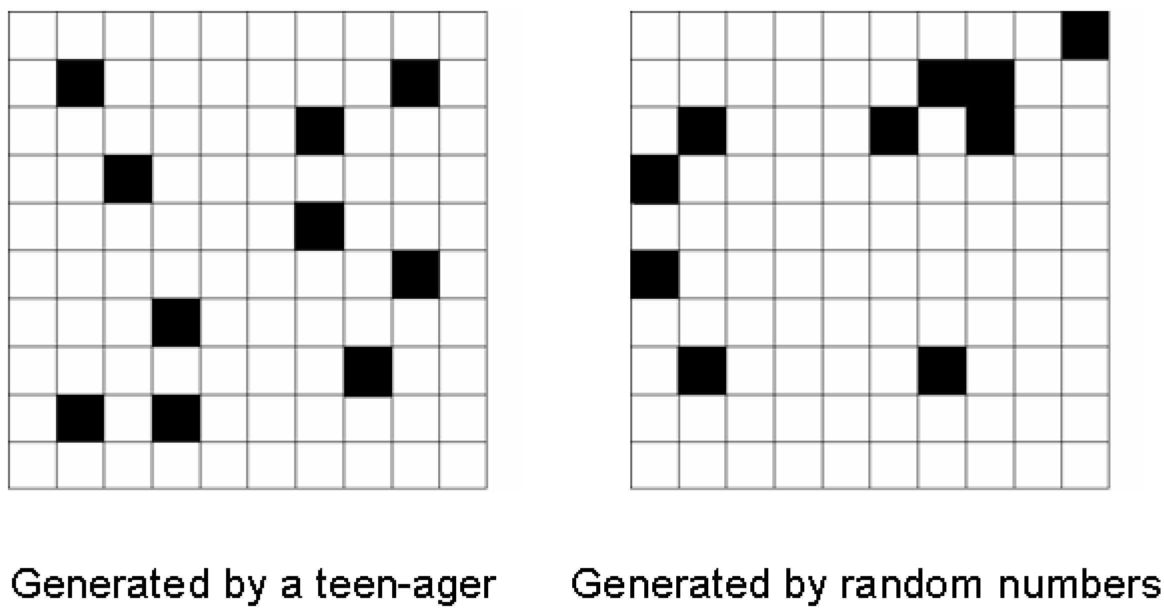


Figure 1

Figure 2. Distributions of hiders' and seekers' choices over the 5×5 grid in the competitive and cooperative conditions in Experiment 1.

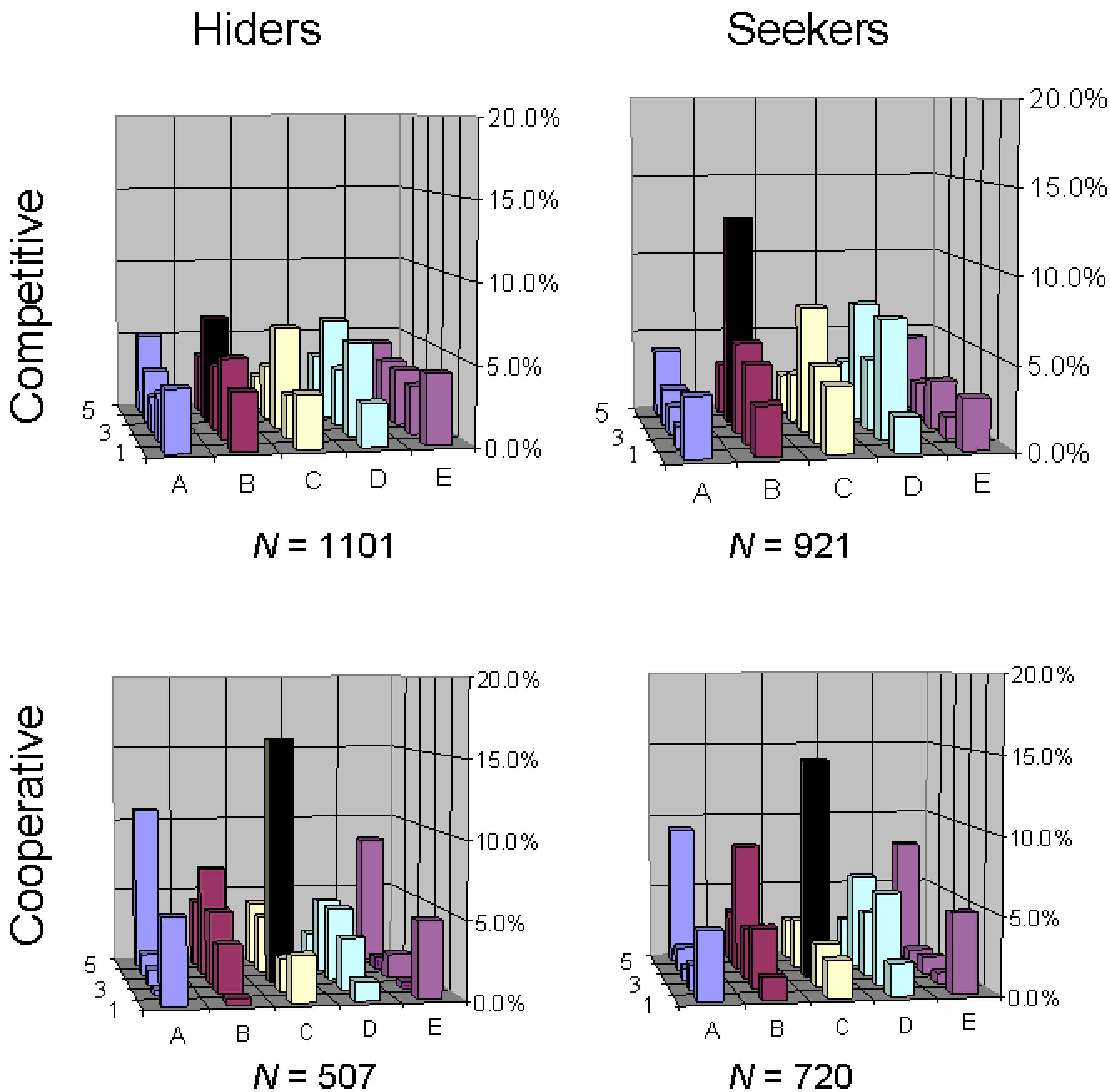

Hiders
Seekers
Competitive
Cooperative
20.0%
15.0%
10.0%
5.0%
0.0%
5
3
1
A
B
C
D
E
N = 1101
N = 921
N = 507
N = 720

Figure 2

Figure 3. Distributions of choices over the 5×5 grid in different conditions in Experiment 1.

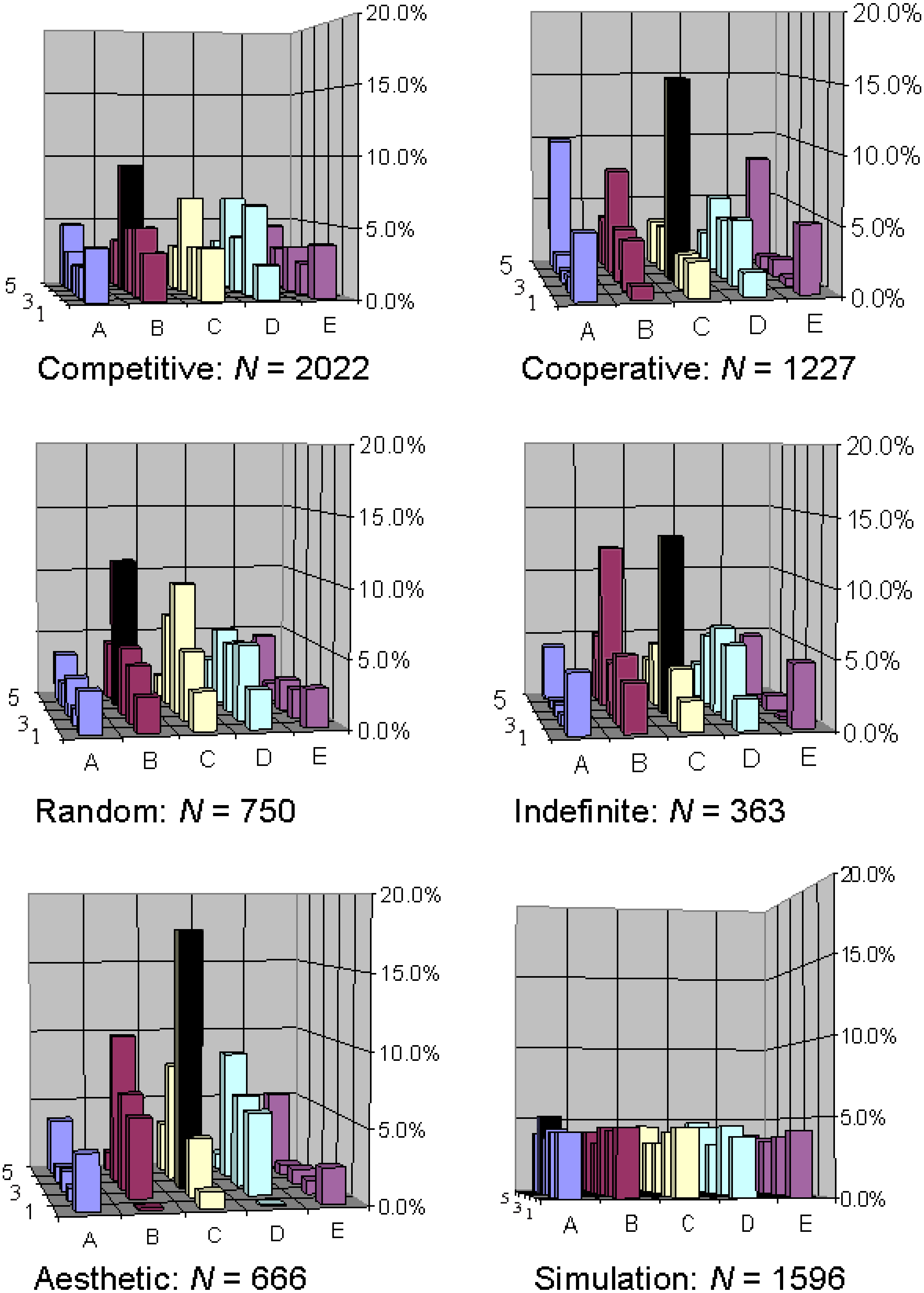

20.0%
15.0%
10.0%
5.0%
0.0%
5
3
1
A
B
C
D
E
Competitive: N = 2022
Cooperative: N = 1227
Random: N = 750
Indefinite: N = 363
Aesthetic: N = 666
Simulation: N = 1596

Figure 3

Figure 4. Six patterns, each selected by over 10 participants, under aesthetic instructions in Experiment 1 (no. of participants = 222).

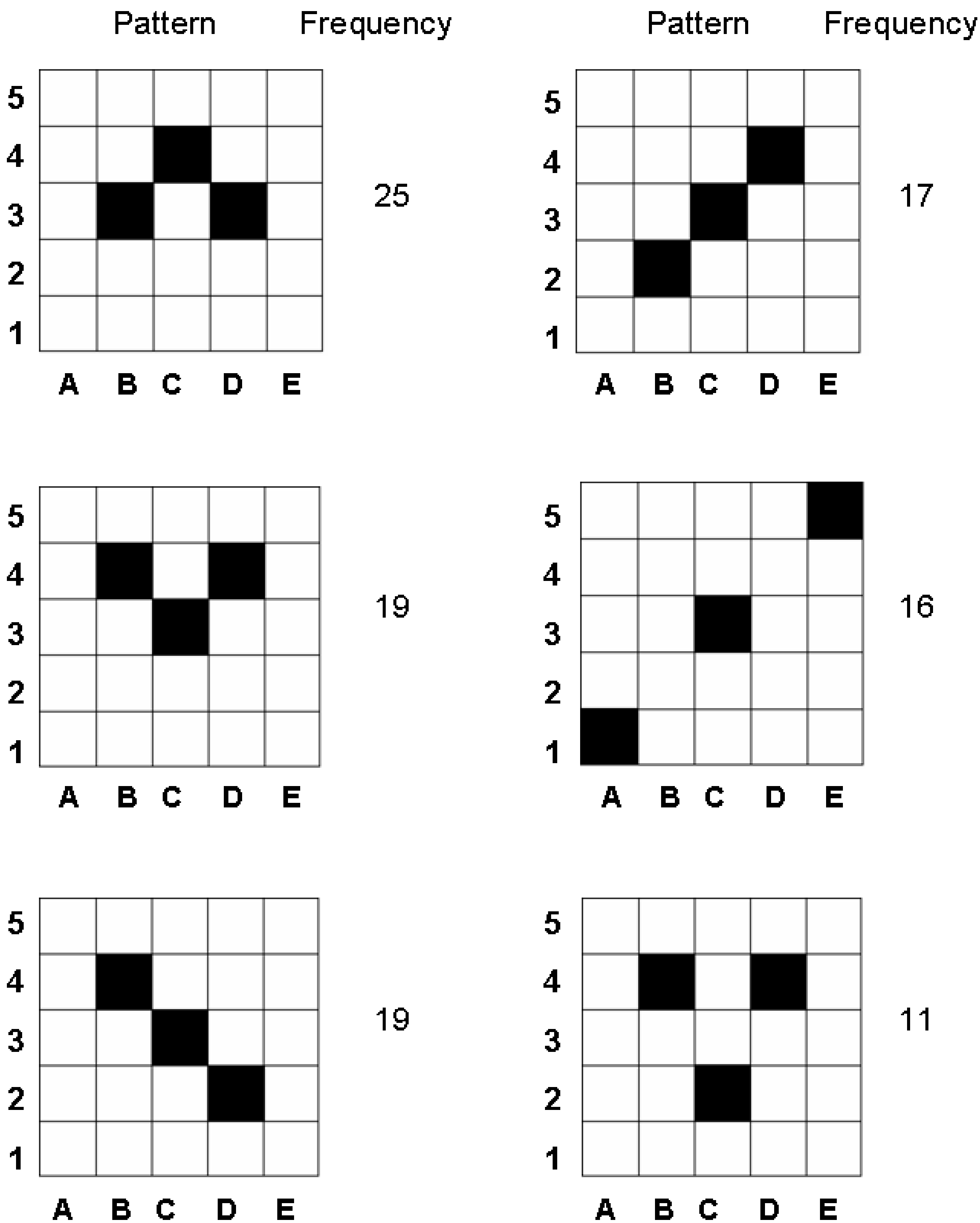

Pattern
Frequency
Pattern
Frequency
5
4
3
2
1
A B C D E
25
17
19
16
19
11

Figure 4

Figure 5. Distributions of hiders' and seekers' choices over the 5×5 grid in the competitive and cooperative conditions in Experiment 2.

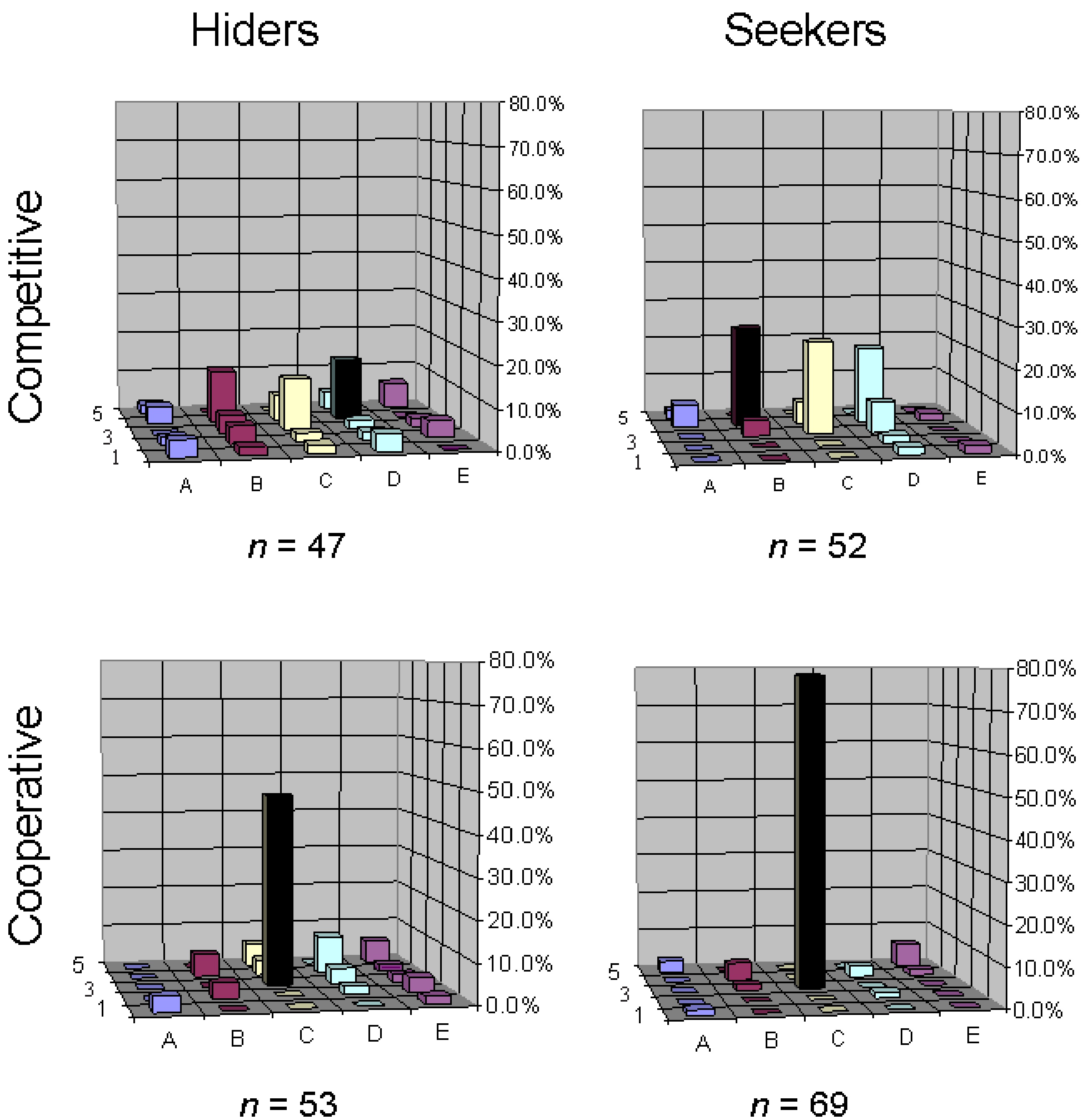

Hiders
Seekers
Competitive
Cooperative
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
5
3
1
A
B
C
D
E
n = 47
n = 52
n = 53
n = 69

Figure 5

Figure 6. Distributions of choices over the 5×5 grid in different conditions in Experiment 2.

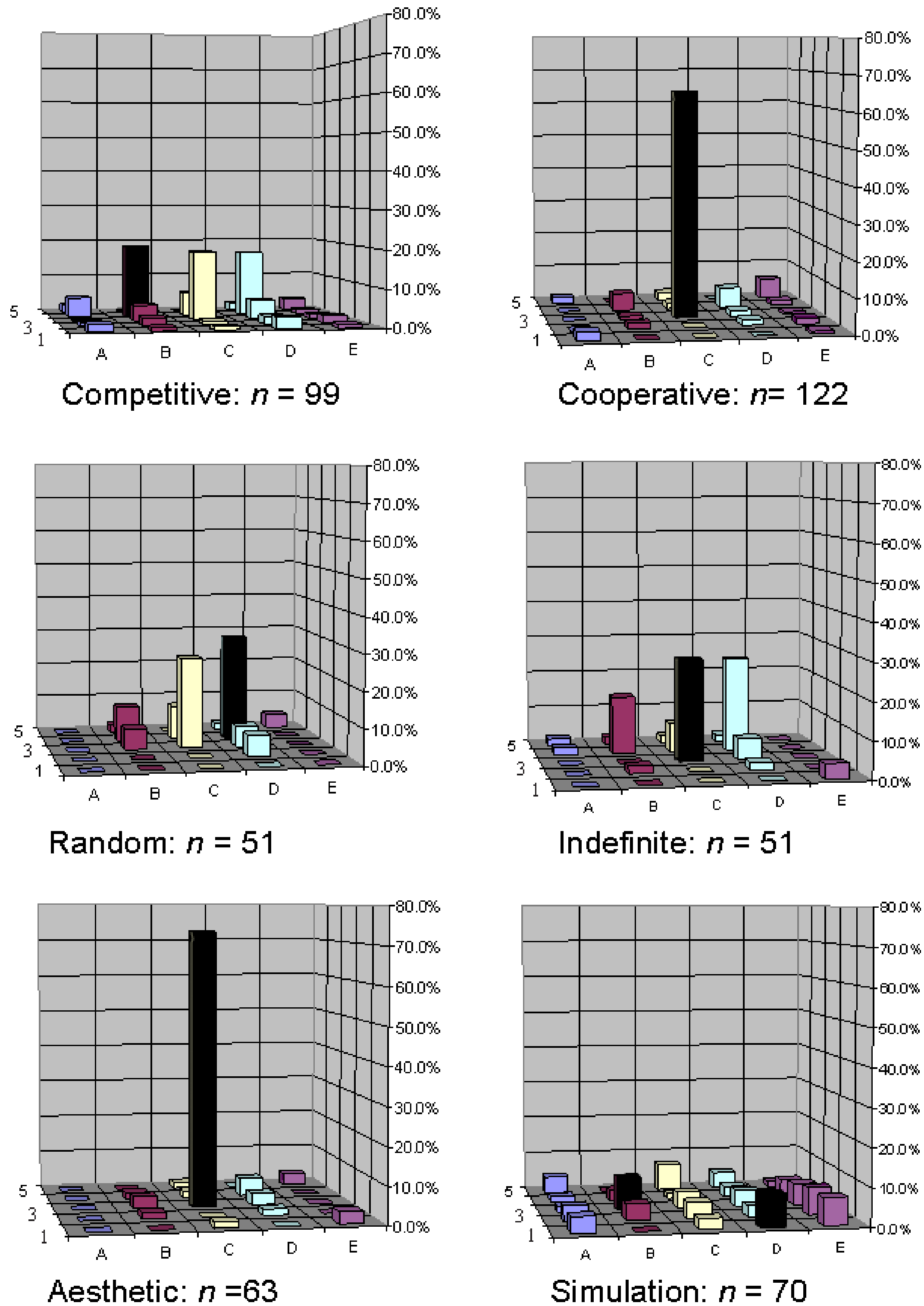
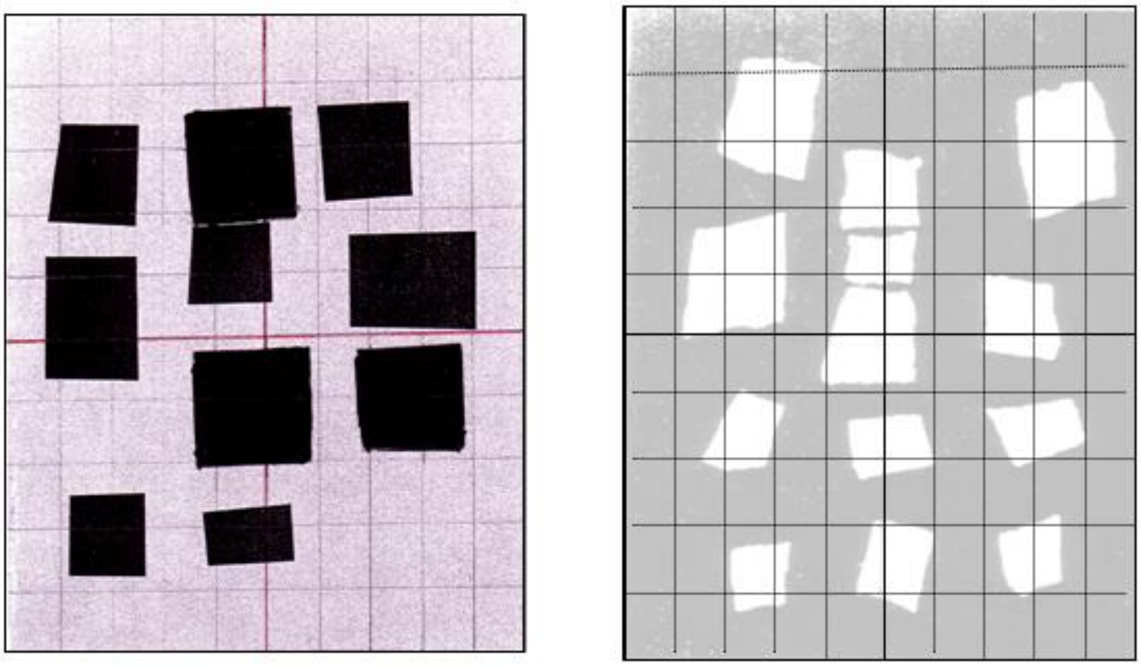

80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
5
3
1
A
B
C
D
E
Competitive: *n* = 99
Cooperative: *n*= 122
Random: *n* = 51
Indefinite: *n* = 51
Aesthetic: *n* =63
Simulation: *n* = 70

Figure 6

<u>Figure 7</u>. Sketches of Arp's "According to the Laws of Chance."

a. 1916 b. 1920



Based on Read (1968, p. 39)

Based on Richter (1964/1965, Illustration 12)

Figure 7