APPLIED COGNITIVE PSYCHOLOGY
*Appl. Cognit. Psychol.* (2008)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/acp.1460

# Pictorial Representations in Statistical Reasoning

GARY L. BRASE*

*Kansas State University, USA*

SUMMARY

In an ongoing debate between two visions of statistical reasoning competency, ecological rationality proponents claim that pictorial representations help tap into the frequency coding mechanisms of the mind, whereas nested sets proponents argue that pictorial representations simply help one to appreciate general subset relationships. Advancing this knowledge into applied areas is hampered by this present disagreement. A series of experiments used Bayesian reasoning problems with different pictorial representations (Venn circles, iconic symbols and Venn circles with dots) to better understand influences on performance across these representation types. Results with various static and interactive presentations of pictures all indicate a consistent advantage for iconic representations. These results are more consistent with an ecological rationality view of how these pictorial representations achieve facilitation in statistical task performance and provide more specific guidance for applied uses. Copyright © 2008 John Wiley & Sons, Ltd.

Certain basic mechanics of how to improve Bayesian reasoning have become clear over the past decade: use frequencies, use them in a nested subset framework and use pictures. For some applied settings, this procedural knowledge is enough. For example, one study found that the majority of counsellors reporting HIV test results provided confusing and inconsistent numerical information, but the researchers described how these problems could be resolved by the adoption of naturally sampled frequencies to communicate results (Gigerenzer, Hoffrage, & Ebert, 1998). To the extent that pictorial representations of numerical information have been experimentally shown to further improve information communication, they should also be useful in applied settings.

Behind these mechanical maxims, however, there is an ongoing debate between two interpretations as to *why* these manipulations have the effects they do, and further understanding at this level can generate more effective applied uses. Proponents of the ecological rationality approach claim that these tactics help tap into the frequency coding mechanisms of the mind, which evolved by natural selection in response to the frequentist nature of objects, events and locations in the natural environment (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). In contrast, others argue that these tactics only help one to appreciate general subset relationships (e.g. Sloman, Over, Slovak, & Stibel, 2003). These two positions disagree in other respects—an important example is whether numbers labelled as 'chances' should be considered probabilities (Girotto & Gonzalez, 2001; Hoffrage, Gigerenzer, Krauss, & Martignon, 2002)—and there is a much more general

*Correspondence to: Gary L. Brase, Department of Psychology, 492 Bluemont Hall, Kansas State University, Manhattan, KS 66506, USA. E-mail: gbrase@ksu.edu

issue of how one conceptualizes human cognitive abilities and standards of rationality (Gigerenzer, Todd, & ABC Research Group, 1999). The focus of the present work, however, is on the specific role of pictorial representations in statistical reasoning, and how this reflects on these two perspectives. Importantly, this issue is also of direct utility to applied contexts. After elaborating on the nature of these two positions, a series of adjudicating experiments are presented.

## Pictures as ecological rationality

One early indication of the usefulness of pictorial representations in statistical reasoning was Gigerenzer and Hoffrage's observation that some study participants used 'pictorial analogs' as a way to successfully achieve Bayesian inference (1995, p. 689). Cosmides and Tooby (1996) added pictures to enhance performance in similar tasks and further developed an 'active pictorial' structure, in which participants were instructed to actively construct the picture. Their rationale was that 'there would be inductive reasoning mechanisms that represent information as frequencies because in our natural environment that is what we would have been encountering: a series of real, discrete, countable events. If true, then the highest levels of Bayesian performance should be elicited when subjects are required to represent the information in the problem as numbers of discrete, countable individuals'. (p. 33). This idea was carried further by Brase, Cosmides, and Tooby (1998), who proposed the *individuation hypothesis*; that an evolved frequency encoding mechanism would be adapted to work best with objects, events or locations that were individuated (i.e. discrete items, rather than aspects or parts of larger items). Finally, much of this research was consolidated and extended by Sedlmeier (1999), who reiterated that the ecological rationality approach 'assumes that pictorial and other external representations are more helpful in solving statistical tasks the more closely they resemble the naturally occurring events they represent'. (p. 65). He also noted that 'participants should. . . show higher solution rates if they experience the sampling process (or a vivid illustration thereof) themselves than if they merely read about it' (p. 39), again based on the idea that pictures approximate actual experienced frequencies. Sedlmeier's work with Bayesian reasoning tasks (1999, chapter 6) found that frequency grids worked better than Venn circles in terms of long-term retention of training on how to complete such tasks. These results suggest that some pictorial representations may be superior to others, but the focus on skill retention over time makes this conclusion less certain.

There is also a long history of findings of quite good simple frequency tracking under naturalistic conditions (see Zacks & Hasher, 2002 for review). For instance, Betsch, Biel, Eddelbuettel, and Mock (1998) found that situations in which base rates were directly experienced led to better use of those base-rates than the same task with information presented as percentages or as frequencies alone, and Hertwig, Barron, Weber, and Erev (2004) found that decisions based on experienced events (using actual behavioural decisions) were better calibrated to the events than mathematically isomorphic judgments based on descriptions (see also Gigerenzer, Hell, & Blank (1988) on active participation in random drawing events versus descriptions of random draws).

## Pictures as illuminations of nested sets

Critics of the ecological rationality perspective have raised objections and alternative explanations. Girotto and Gonzalez (2001) cited work (i.e. Gluck & Bower, 1988; Shanks,



DOI: 10.1002/acp

1990; Nosofsky, Kruschke, & McKinley, 1992) that argues against an experience-based competency in frequency tracking. They suggest that 'if human competence in frequency information processing had emerged through natural selection, reasoners would perform better on problems closer to the natural settings in which information is acquired than on the more artificial verbal problems. In fact, reasoners make biased frequency predictions in both types of problems' (p. 271).

Also in contrast to the ecological rationality account, Yamagishi (2003) proposed a roulette-wheel diagram as, 'a nested-sets instruction that does not necessarily call for attention to frequency... [yet] shows the relationship between the prior and the posterior probabilities'. (p. 98). Yamagishi argued that 'the graphical nature of [roulette-wheel figures] take advantage of people's automatic visual computation in grasping the relationship between the prior and posterior probabilities'. (p. 105). This is not just a difference in emphasis or semantics, according to this proposal; Yamagishi argued that improved performance with the use of this representation was due to the clarity of the nested-set presentation and not due to any use of frequencies (i.e. discrete, countable items). In a similar fashion, Sloman et al. (2003) proposed that 'representing instances can reveal the set structure of a problem. ... Most representational schemes that identify instances and the categories they belong to will automatically also specify the set structures relating the categories (e.g. Euler circles, mental models, taxonomic hierarchies)' (p. 298).

With impressive findings in support of both human competence and incompetence in frequency information processing, there is a legitimate discrepancy in the literature. A conciliatory position might propose that good human frequency tracking is constrained to certain domains, which would be sensible because tracking of all conceivable frequencies in the environment would outstrip even the most optimistic estimates of human computational abilities. For example, the individuation hypothesis (Brase et al., 1998) proposed that frequency tracking is more intuitively tuned to tracking whole objects, events and locations (as opposed to aspects or parts of these items), thereby achieving more computational tractability. Even this proposal, however, serves to highlight the need to understand more clearly the nature of this domain-specificity in frequency tracking (and hence is contrary to the nested sets position).

## 'Chances' as a testing ground

A basic methodological problem in this area is that instantiations of natural sampling frameworks tend to be frequentist in nature (e.g. using counts of whole numbers, which certainly do not conform to the normalization between 0 and 1 for single-event probabilities), yet what is needed experimentally is a separation of these two factors (i.e. existence/absence of natural sampling representation and type of numerical representation). Girotto and Gonzalez (2001) argued that information presented as 'chances' are probabilities—and not frequencies—that can nevertheless be put into nested set relationships (see also Sloman et al., 2003, who concur). This position has been criticized, however, on the grounds that it is an artificial and arbitrary labelling of information as probabilities rather than frequencies (Brase, 2002; Hoffrage et al., 2002), based primarily on the researchers' interpretation which may not be consistent with interpretations by research participants. In fact, Brase (2008) found that the use of whole number 'chances' is sufficiently ambiguous to elicit either probabilistic or frequentist representations in individuals, as suggested by these critics. (Generally, Brase (2008) found, in follow-up questions to Bayesian reasoning tasks, that there was a 60/30 split in how participants



DOI: 10.1002/acp

described their representations; chances were identified as single-event probabilities by about 60% of people, as frequencies by 30%, and as something 'other' than those options by about 10% of people. Frequencies showed a similar—but opposite—60/30 division in interpretation.)

## EXPERIMENT 1

What has become clear is that there is some ambiguity in the interpretation of information expressed as 'chances' (manifested both as disagreement at the theoretical level, and as different interpretations at the level of individual research participants). This situation provides a unique route by which natural sampling frameworks (including the use of whole numbers) can be created which are nevertheless interpretable as either frequencies or as single-event probabilities. Such an ambiguous situation can be useful here as a context in which participants may be swayed one way (frequency interpretation) or the other (probabilistic interpretation) by relatively subtle changes in context. Specifically, given the same style and format of problem text, Bayesian reasoning may be significantly influenced by the use of pictorial representations, but contrasting theoretical accounts lead to slightly different hypotheses about which pictures should facilitate of Bayesian reasoning.

(a) An approach focused on just the appreciation of nested-set relationships would predict that any pictorial representation (e.g. icons, Venn diagrams or roulette wheel diagrams) would enhance the perception of the nested-set relations that exist within the problem, and thereby improve Bayesian reasoning.
(b) In contrast, an ecological rationality approach predicts that pictorial representations which utilize individuated entities that represent specific object, events or locations (e.g. discrete icons) will elicit frequentist representations (and therefore improve Bayesian reasoning) better than non-individuated entities because these would effectively tap into intuitive cognitive mechanisms for tracking frequencies of real-world object, events or locations.
(c) Performance based on a pictorial representation that combines both Venn circles and something akin to icons (i.e. dots within Venn circles) can be quite informative. If the cognitive process in question tends to respond best to experiences of clearly individuated items, then performance should suffer some in comparison to the more clearly discrete icons but not fall to the level of empty Venn circles. If the cognitive process in question tends to respond best to experiences of clear nested-sets (and individuated, frequentist-type elements are irrelevant), then performance should be as good as with the empty Venn circles (which would presumably be very good).

### Method

*Participants*

The participants were 412 undergraduates (187 males, 225 females) from a large Midwestern university. All participants completed the study as partial fulfilment of an introductory psychology class requirement. The average age of participants was 18.7 years.

*Materials and procedure*

After reading and signing a separate informed consent form, each participant was given a single sheet with some brief instructions, a Bayesian reasoning problem and a follow-up



DOI: 10.1002/acp

question on the reverse side of the sheet. The Bayesian reasoning problem was a version of the 'medical diagnosis problem' (Cosmides and Tooby, 1996), with an accompanying pictorial representation for the task. The Bayesian reasoning task used in these studies differs from those used by Sloman et al. (2003) and others in some minor—but important—respects. Whereas some previous studies have used a Bayesian reasoning task with limiting-case situations (e.g. '1 out of 1000'), this type of scenario can lead to shortcut solutions that are adaptive but not truly Bayesian inferences (see Gigerenzer & Hoffrage, 1995; Macchi & Mosconi, 1998). This potential confusion is exacerbated if lenient criteria are applied in judging correct responses (as in Sloman et al., 2003).

In this and the following studies, participant instructions and Bayesian problems were consistent across all conditions. The instructions told participants to answer the problem and to report the 'typical' outcome if they believed that the answer may change each time the situation described occurs. The actual Bayesian problem was as follows:

> There is a newly discovered disease, Disease *X*, which is transmitted by a bacterial infection. Here is some information about the current research on Disease *X* and efforts to test for the infection that causes it.
>
> A person has 6 chances out of 100 of having the infection. There is a test to detect whether or not a person has this infection, but it is not perfect. Specifically, only 4 of the 6 chances of having the infection were associated with a positive reaction from the test. On the other hand, 16 of the remaining 94 chances of not having the infection (that is, being perfectly healthy) were also associated with a positive reaction from the test.

This information was followed by one of four conditions for pictorial representation of the problem, into which participants were randomly assigned: (1) a control problem with no pictorial representation, (2) a problem restatement using Venn circles, (3) a problem restatement using Venn circles with large and visible dots within each area of the circles and (4) a problem restatement using icons (see Figure 1). Below the pictorial representation was the Bayesian inference question:

> Imagine Michael is tested now. Out of a total of 100 chances, Michael has _____ chance(s) of positive reaction from the test, _____ of which will be associated with actually having the infection.

A question following the Bayesian task (on the opposite side of the paper) asked if the participant 'had any training or education in how to combine conditional probabilities (tasks such as this)' and if so, to detail that experience. Thirteen participants gave affirmative responses to this question and were able to identify relevant training (e.g. in Bayesian reasoning, probability or posterior odds, etc.), so they were excluded from the results and analysis. As a result, there were 96 remaining participants in the no picture condition, 98 participants in the Venn condition, 108 participants in the dotted Venn condition and 95 participants in the icons condition.

## Results

As shown in Table 1, participants were much more likely to correctly answer this task (i.e. 20 chances of a positive reaction, 4 of which are associated with infection) when the Bayesian problem was accompanied by iconic representations, as compared to either no



DOI: 10.1002/acp

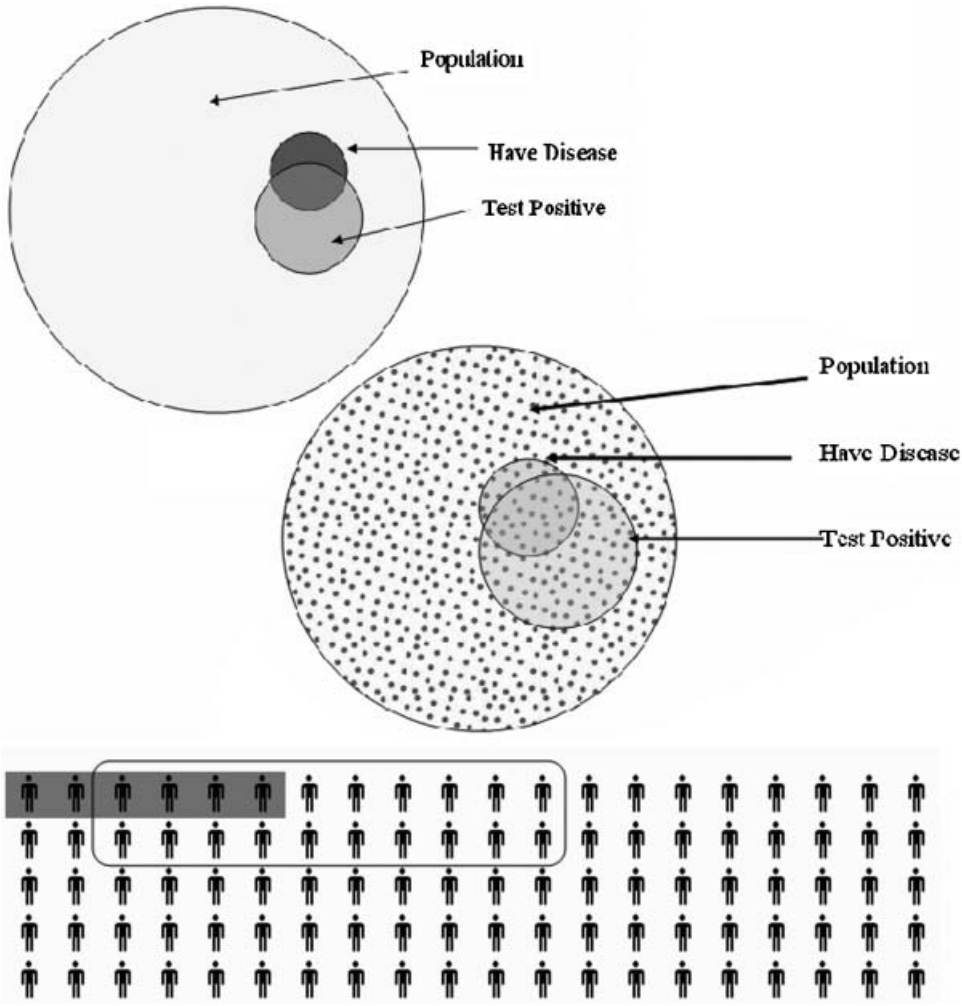


Figure 1. Stimuli used in Experiment 1; Pictorial representations used to depict the Bayesian reasoning problem. The Venn diagrams (top two) were preceded by the following written description: The picture below summarizes the above information, and is provided for you to use in the process of answering the following item. All the chances in this situation are represented by the largest circle. The smallest circle represents the chances of having the infection. The slightly larger, overlapping circle represents the chances of a positive reaction to the test The icon representation (bottom) was preceded by the following written description: The picture below summarizes the above information, and is provided for you to use in the process of answering the following item. 100 chances are represented by the 100 figures printed below (in 5 rows of 20). Figures that are darkened are those chances with the infection. Figures that are circled are those chances who have a positive reaction to the test

picture aid (48.4% vs. 35.4%: $z = 1.82$, $p = .035$, $p_{rep} = .90$, $\eta = .26$.) or Venn circles (48.4% vs. 34.7%: $z = 1.93$, $p = .027$, $p_{rep} = .91$, $\eta = .28$; using difference of proportions tests). Performance on the problem when accompanied by the dotted Venn representation (41.7%) was intermediate between the plain Venn circles and the icons, as predicted by the ecological rationality approach. Specifically, the dotted Venn condition was about equidistant from the plain Venn and icon conditions, and was not statistically significant



DOI: 10.1002/acp

Table 1. Percentages of participants in Experiment 1 reaching the correct Bayesian response, by type of pictorial representation provided

| Type of pictorial representation | Overall correct (%) |
|---|---|
| No picture ($n = 96$) | 35.4 |
| Venn ($n = 98$) | 34.7 |
| Dotted Venn ($n = 108$) | 41.7 |
| Icons ($n = 95$) | 48.4 |

*Note*: In this experiment, and in all subsequent experiments, the correct answer was either the modal response or close to the modal response. More detailed descriptions of the frequencies with which different incorrect responses were made (or the raw data) are available from the author.

from either (34.7% vs. 41.7% and 41.7% vs. 48.4%: $z = 1.03$, $p = .15$, $p_{rep} = .77$, $\eta = .14$ and: $z = 0.96$, $p = .17$, $p_{rep} = .75$, $\eta = .13$, respectively). Finally, the use of empty Venn circles showed no effect, as compared to no picture aid (34.7% vs. 35.4%: $z = 0.10$, $p = .46$, $p_{rep} = .53$, $\eta = .01$).

This finding is remarkable in that only the pictorial representation was manipulated, whereas everything else—the text of the task, the format of the numbers and the natural sampling structure (a.k.a., subset relations) across the numbers—was held constant across the conditions. Nevertheless, there was a measurable effect on performance when icons were used to represent the task situation, as compared to both no pictures and Venn circle representation.

## EXPERIMENT 2

The results of Experiment 1 indicate that pictorial representations which more closely mimic the individuated nature of natural frequency derived information elicit better Bayesian reasoning. One could argue, however, that the presentation of icon representations allows for a strategy of actually counting the icons to obtain the answer, thus providing an additional route to the correct response. If this were the case, then including instructions to actively count out the answer—using a provided representational format—should eliminate this unfair advantage and 'level the field' across the different formats. Indeed, Cosmides and Tooby (1996) found that active involvement with a pictorial representation (of icons) boosted performance. If the facilitation is due to the use of individuated icons *per se,* however, the active involvement with a pictorial representation that is not supportive of an individuated, frequency representation should not necessarily show the same boost in performance.

### Participants

The participants were 149 undergraduates (56 males, 93 females) from a large, Midwestern university. Nine participants were excluded from the results either for incomplete responses or after indicating that they had relevant training in Bayesian or probability tasks. All participants completed the study as partial fulfilment of an introductory psychology class requirement. The average age of participants was 18.7 years.



DOI: 10.1002/acp

## Materials and procedure

The materials and procedure were the same as in Experiment 1, utilizing all three pictorial representation conditions developed previously. Each of these representations was modified, though, to be 'active' representations in which the participants were instructed to complete the picture so that it corresponded to the information in the Bayesian inference task. Specifically, participants began with either: (a) a field of 100 icons as in the bottom of Figure 1, but without any shading or circled items, (b) a single empty Venn circle or (c) a single dot-filled Venn circle (again, as in Figure 1, but with the inner circles removed). Participants were instructed in each condition that the picture represented 'all the chances in this situation', and to either cross out figures (for icons) or create a circle (for Venn circles) to represent chances of having an infection. The next instruction was to make a circle to represent the chances of a positive reaction to the test. Below the pictorial representation area was the Bayesian inference question.

## Results

Participants were again more likely to reach the correct Bayesian answer when the problem was accompanied by iconic representations, as compared to either empty or dot-filled Venn circles (49.0% vs. 30.0%: $z = 1.93$, $p = .028$, $p_{rep} = .91$, $\eta = .39$ and 49.0% vs. 28.0%: $z = 2.15$, $p = .017$, $p_{rep} = .93$, $\eta = .44$ see Table 2). The lack of difference between empty Venn circles and dot-filled Venn circles does not support the idea that participants gained particular advantage by having items to count as a solution strategy. Many participants were clearly meticulous in developing accurate (if convoluted) circles to enclose the correct numbers of dots in their Venn circles, yet this did not improve performance.

It is worth noting that participants in the active pictorial conditions (Experiment 2) did not perform better than analogous pictorial conditions that did not require active construction (Experiment 1), which runs counter to previous facilitative effects of active pictorial construction (Cosmides & Tooby, 1996). One potential way to reconcile these findings is to note that many of the active constructions of pictorial representations were not correctly completed, and incorrect constructions could just as easily have confused participants rather than helped them. In fact, if one looks at participants who created correct pictorial constructions versus those who created incorrect constructions (Table 2), this does appear to be an important factor. Interestingly, correct

Table 2. Percentages of participants in Experiment 2 reaching the correct Bayesian response, by type of pictorial representation provided and participants' pictorial construction

| | Level of correct performance | | |
|---|---|---|---|
| Type of pictorial representation | Overall (%) | Correct construction (%) | Incorrect construction (%) |
| Active Venn ($n = 50$) | 30.0 | 71.4 ($n = 7$) | 23.3 ($n = 43$) |
| Active dotted Venn ($n = 50$) | 28.0 | 66.7 ($n = 3$) | 25.5 ($n = 47$) |
| Active icons ($n = 49$) | 49.0 | 66.7 ($n = 27$) | 27.3 ($n = 22$) |



DOI: 10.1002/acp

constructions were much more common in the icon representation condition, and across all conditions the type of pictorial construction (correct or incorrect) was not perfectly associated with correct or incorrect numerical answers.

## EXPERIMENT 3

Experiments 1 and 2 both found an advantage of iconic representations over Venn representation for facilitating Bayesian reasoning. These previous experiments, however, are still subject to a couple alternative explanations. One purpose of Experiment 3 was to clarify the possibility that people could be counting and using the dots within the dotted Venn representations while simultaneously ignoring all of the problem text. In such a case, the original dotted Venn representation could produce contradictory answers (the original dotted Venn is an even dot pattern superimposed over the Venn circles, and therefore has some dots that are intersected by or touching lines and has numbers of dots within some circles that do not correspond to the problem text). To resolve this issue a new dotted Venn representation was created that used 100 dots, individually placed within the Venn circles and in locations that correspond exactly and unambiguously to the problem text. A second purpose of Experiment 3 was to address a concern that the information representations using icons have—in previous experiments—been neatly organized to provide a clean and clear structure; categories of situations have been put together in blocks or circles. One could argue that this implicitly creates a Venn-like representation within the icon representation conditions, constituting an 'unfair advantage'. To resolve this issue, a new icon representation was created that distributed the instances of different situations throughout the representation.

### Participants

The participants were 166 undergraduates (65 males, 101 females) from a large, Midwestern university. One participant was excluded from the results for incomplete responses. All participants completed the study as partial fulfilment of an introductory psychology class requirement. The average age of participants was 18.8 years.

### Materials and procedure

The materials and procedure were the same as in Experiments 1 and 2, utilizing four pictorial representation conditions: (a) a Venn representation with open fields (the same as Experiment 1), (b) a new Venn representation with clarified and completely accurate dots within it (see Figure 2), (c) an icon representation with information organized (the same as Experiment 1) and (d) an icon representation with information randomly distributed (see Figure 2).

### Results

There was no difference in performance based on whether the information in iconic representations were grouped (47.6%) or spaced out (47.6%; see Table 3). Once again, however, participants were again more likely to reach the correct Bayesian answer when



DOI: 10.1002/acp

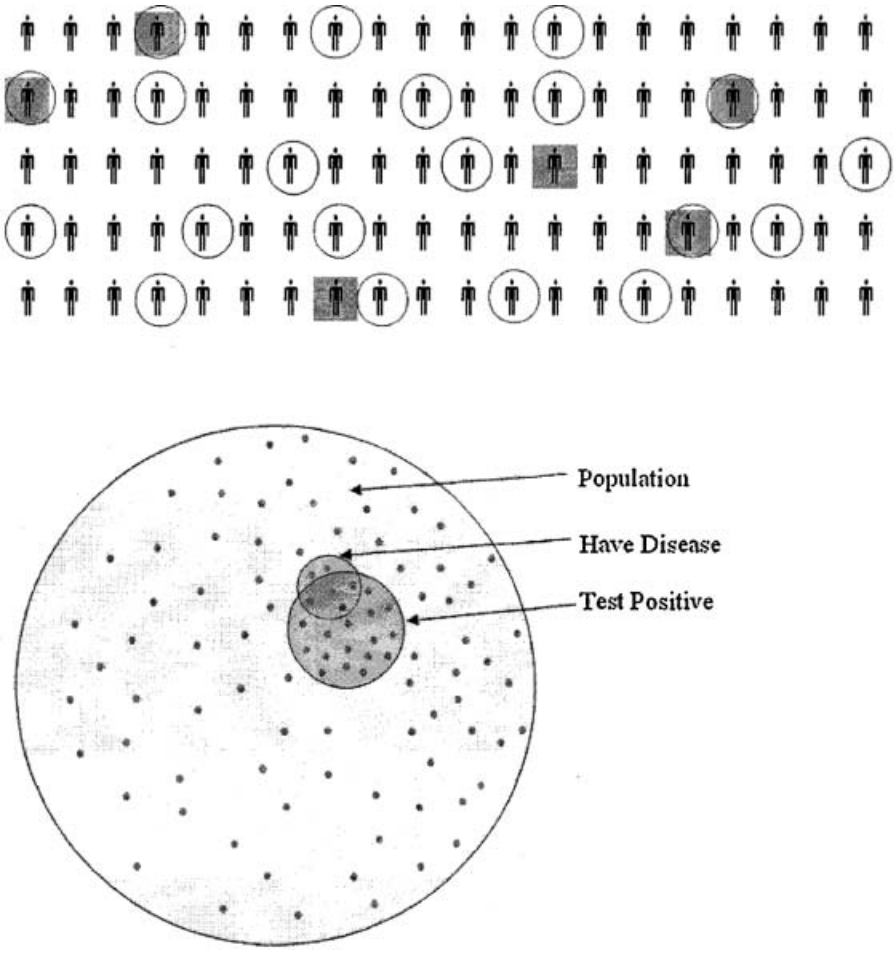


Figure 2. Stimuli used in Experiment 3 included replications of empty Venn picture and icons pictures used in Experiment 1, and the pictures shown here: A randomly spaced events icon picture and a Venn picture with exactly 100 dots in completely accurate locations

the problem was accompanied by iconic representations, as compared to either empty or dot-filled Venn circles (47.6% vs. 26.8%: $z = 1.96$, $p = .027$, $p_{rep} = .91$, $\eta = .43$ and 47.6% vs. 20.0%: $z = 2.64$, $p = .005$, $p_{rep} = .97$, $\eta = .60$; see Table 2). These results also replicate the lack of difference between the empty Venn and dotted Venn representation conditions (20.0% vs. 26.8%: $z = 0.73$, $p = .235$, $p_{rep} = .695$, $\eta = .16$).

Table 3. Percentages of participants in Experiment 3 reaching the correct Bayesian response, by type of pictorial representation provided

| Type of pictorial representation | Overall correct (%) |
|---|---|
| Icons, grouped ($n = 42$) | 47.6 |
| Icons, individual ($n = 42$) | 47.6 |
| Venn with dots ($n = 41$) | 26.8 |
| Venn with open fields ($n = 40$) | 20.0 |



DOI: 10.1002/acp

## DISCUSSION

Across a series of experiments, it was consistently found that people do significantly better at Bayesian reasoning tasks that are accompanied by iconic pictorial representations, as compared to either pictorial representations in the form of continuous fields (Venn circles) or no pictorial representation at all. Adding dots to Venn circles may sometimes help performance, but it does not appear to be sufficient to generate facilitation similar to that found with iconic representations. Active construction of pictorial representations using these different formats found the same pattern of superior performance with iconic representations. Pictorial representations that used randomly spaced information throughout the icon representation were just as effective as when classes of entities were grouped within the icons. Finally, this result is not due to the ability to simply count icons because: (a) Venn representations filled dots can similarly be counted (yet do not show a similar effect), (b) specific instructions to actively construct the correct representation are only facilitative for iconic representations and (c) there was no difference in overall performance with iconic representations when respondents were instructed to count (Exp. 2, 49.0%) versus when they were not instructed to count (Exp. 1, 48.4%).

A few aspects of these results deserve highlighting. First, all these effects were obtained while using one, identical Bayesian reasoning problem. Thus, the differences in performance across conditions resulted from just the post-text pictorial representations, and significant differences therefore carry particular weight in that the conditions were in all other respects identical. On the other hand, the strict experimental control achieved by using the same reasoning problem across all these studies does raise a need for further work to establish the generalizability of these results; will the same effects of different pictorial representations be found in problems involving other types of contexts. For that matter, the effects of pictorial representations also remain to be evaluated for mathematical problems beyond Bayesian reasoning situations.

A second part to highlight is the irony that, as a general rule, Venn circles are excellently suited for revealing the nested-set structure of Bayesian tasks (Sloman et al., 2003; Yamagishi, 2003); they provide a clarity and simplicity of display that 100 individual icons will never have. Yet Venn circles did not facilitate Bayesian reasoning in any of these studies to the same extent as an iconic representation. Indeed, one could have argued that the circled and highlighted icons used in Experiments 1 and 2 constituted Venn-like structures that were contributing to the performance in these conditions (although it would be unclear, in that case, why participants in these conditions were performing *better* than actual empty Venn circles). That argument is moot, though, because Experiment 3 demonstrated that better performance was associated with the use of icons with no nested-set structures at all. Overall, these results indicate that iconic representational conditions, which better approximate actual ecological presentations, can facilitate correct Bayesian reasoning, as predicted by the ecological rationality approach.

A final aspect of the results that deserves specific discussion is the relationship between 'dots', as used in some of the Venn circles, and the icons used in other conditions. What makes a dot less of an 'icon' than the items labelled here as icons? That is, can't the dots be used to represent specific, individuated items just as well as other pattern of ink on a page? This is certainly a question that can be investigated empirically, and two considerations are likely to be important in such research. First, the icons used in this research (little human-like figures) were selected to be easy to represent as individual people and they were arranged in rows to de-emphasize any subset-like grouping. Thus, it is possible that



DOI: 10.1002/acp

this type of icon was perceived as more ecologically realistic (than dots, at least) in terms of representing individuals. Second, it theoretically should be possible to improve participant performance with dots as stimuli by emphasizing that the dots are representative of discrete items (e.g. by using explicit instructions to that effect). In some respects, then, one can think of the different types of pictures and associated instructions as existing along a continuum between totally abstract set-relations (e.g. pure Venn circles occupying conceptual space) and totally concrete stimuli (e.g. pictures or video that mimics actual experience). The present studies have sampled a few locations along this continuum that are important for theoretical reasons and for their potential applied utility.

In practical terms, this research indicates that not all types of pictorial representations are equivalently useful. Representations that better approximate natural sampling of frequencies (occurrences of individuated objects, events or locations) tend to elicit better Bayesian reasoning. Even if such sampling is done in more abstract terms (e.g. using icons on a page that represent objects, events or locations) rather than actual experience, such natural sampling experiences can be valuable in applied settings.

What would pictorial representations of numerical information look like in applied settings? Think about a person visiting a doctor to discuss a recent test result (e.g. an HIV test as discussed in the introduction, a prenatal test, or a cancer test). Even though the patient may not have a highly developed mathematics education, he or she must be given the test results—including the base rate, false positive rate and false negative rate—and is the person ultimately responsible for determining the appropriate response to those test results. Presenting numerical information using naturally sampled frequencies, supplemented with iconic representations, will almost certainly facilitate this comprehension process. Indeed, even doctors' comprehension of information can be improved by the use of these presentation strategies (Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2000).

Another example of an applied use of this work is in the area of mathematics education. Some research has already begun, based on principles of ecological rationality, that uses physical items to help young children acquire mathematical concepts such as sampling, proportions and conditional likelihood (Kurz-Milcke & Martignon, 2006; Martignon, Laskey & Kurz-Milcke, 2007; see also Zhu & Gigerenzer, 2006). Similar advantages of these presentation formats can be attained in presenting information in legal settings and in evaluating information relevant to terrorism-related assessments (e.g. the base-rate of ethnic and religious group memberships, relative to identifying terrorism suspects).

## ACKNOWLEDGEMENTS

The author thanks Angela Zellmer for valuable assistance in conducting the research, and Robyn Dawes, Mike Dougherty, Laura Martingnon, Sandra Brase and several anonymous reviewers for their comments and advice regarding this research.

## REFERENCES

Betsch, T., Biel, G. M., Eddelbuettel, C., & Mock, A. (1998). Natural sampling and base-rate neglect. *European Journal of Social Psychology*, *28*, 269–273.



DOI: 10.1002/acp

Brase, G. L. (2002). Ecological and evolutionary validity: Comments on Johnson-Laird, Legrenzi, Girotto, Legrenzi, & Caverni's (1999) mental model theory of extensional reasoning. *Psychological Review*, *109*, 722–728.

Brase, G. L. (2008). Frequency interpretation of ambiguous statistical information facilitates Bayesian reasoning. *Psychonomic Bulletin and Review*, *15*, 284–289.

Brase, G. L., Cosmides, L., & Tooby, J. (1998). Individuation, counting and statistical inference: The role of frequency and whole-object representations in judgment under uncertainty. *Journal of Experimental Psychology: General*, *127*, 3–21.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*, 1–73.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684–704.

Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 513–525.

Gigerenzer, G., Hoffrage, U., & Ebert, A. (1998). AIDS counselling for low-risk clients. *AIDS Care*, *10*, 197–211.

Gigerenzer, G., & Todd, P. M. the ABC Research Group. (1999). *Simple heuristics that make us smart*. Oxford, UK: Oxford University Press.

Girotto, V., & Gonzalez, M. (2001). Solving probabilistic and statistical problems: A matter of information structure and question form. *Cognition*, *78*, 247–276.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*, 227–247.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*, 534–539.

Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science*, *290*, 2261–2262.

Hoffrage, U., Gigerenzer, G., Krauss, S., & Martignon, L. (2002). Representation facilitates reasoning: What natural frequencies are and what they are not. *Cognition*, *84*, 343–352.

Kurz-Milcke, E., & Martignon, L. (2006). Lebendige Urnen und ereignisreiche Bäume: Überlegungen und Versuche zu einer Didaktik der Stochastik in Der Grundschule. In J. Meyer (Ed.), *Anregungen zum Stochastikunterricht, Bd.3* (pp. 181–203). Hildesheim: Franzbecker.

Macchi, L., & Mosconi, G. (1998). Computational features vs frequentist phrasing in the base-rate fallacy. *Swiss Journal of Psychology/Schweizerische Zeitschrift Fuer Psychologie Revue Suisse De Psychologie*, *57*, 79–85.

Martignon, L., Laskey, K., & Kurz-Milcke, E. (2007). Transparent urns and coloured tinker cubes for natural stochastics in primary school. *Proceedings of the Cerme 5 (Fifth Conference of the European Reunion of Mathematics Educators)*, 752–780.

Nosofsky, R. M., Kruschke, J. K., & McKinley, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 211–233.

Sedlmeier, P. (1999). *Improving statistical reasoning: Theoretical models and practical implications*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., Publishers.

Shanks, D. R. (1990). Connectionism and the learning of probabilistic concepts. *Quarterly Journal of Experimental Psychology*, *42A*, 209–237.

Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes*, *91*, 296–309.

Yamagishi, K. (2003). Facilitating normative judgments of conditional probability: Frequency or nested sets? *Experimental Psychology*, *50*, 97–106.

Zacks, R. T., & Hasher, L. (2002). Frequency processing: A twenty-five year perspective. In P. Sedlmeier (Ed.), *ETC frequency processing and cognition* (pp. 21–36). London: Oxford University Press.

Zhu, L., & Gigerenzer, G. (2006). Children can solve Bayesian problems: The role of representation in mental computation. *Cognition*, *98*, 287–308.



DOI: 10.1002/acp