Cambridge University Press & Assessment
978-1-316-51865-6 — Sampling in Judgment and Decision Making
Edited by Klaus Fiedler , Peter Juslin , Jerker Denrell
Excerpt
More Information

PART I

Historical Review of Sampling Perspectives and Major Paradigms

CHAPTER 1

The Theoretical Beauty and Fertility of Sampling Approaches
A Historical and Meta-Theoretical Review

Klaus Fiedler, Peter Juslin and Jerker Denrell

1.1 Introduction

The topic of the present volume, sampling approaches to judgment and decision-making (JDM) research, is ideally suited to illustrate the power and fertility of theory-driven research and theorizing in a flourishing area of behavioral science. The last two decades of rationality research, in psychology, economics, philosophy, biology, and computer science, are replete with ideas borrowed from statistical sampling models that place distinct constraints on information transition processes. These sampling approaches highlight the wisdom gained from Kurt Lewin and Egon Brunswik that in order to understand cognitive and motivational processes within the individual, it is first of all essential to understand the structure and distribution of the environmental stimulus input that impinges on the individual's mind. This is exactly the focus of sampling-theory approaches.

The environmental input triggers, enables, constrains, and biases the information transmission process before any cognitive processes come into play. Because the information offered in newspapers, TV, Internet, textbooks, and literature, or through personal communication is hardly ever an unbiased representative sample of the world, but is inevitably selective and biased toward some and against other topics and sources, a comprehensive theory of judgment and decision-making must take the ecology into account. Importantly, the information input is not only reflective of existing biases of a wicked environment. It is also empowered by the statistical strength and reliability of a distributed array of observations, the statistical properties of which are well understood. So, the challenges of a potentially biased "wicked" environment (Hogarth et al., 2015) come

Work on this chapter was supported to grants awarded to the first author by the Deutsche Forschungsgemeinschaft (Fi 294 / 29 and 30) and to the second author by the Marcus and Amalia Wallenberg Foundation (MAW 2016.0132).

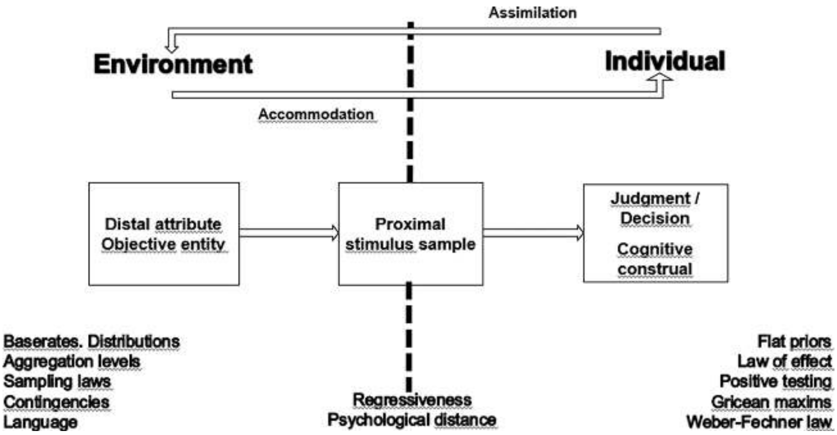

Figure 1.1 Two stages of information transmission from a cognitive-ecological perspective
(Fiedler, 2007; Fiedler & Kutzner, 2015; Fiedler & Wänke, 2009)

along with normative instruments for debiasing and separating the wheat from the chaff.

For a comprehensive theory of judgments and decisions in a probabilistic world, the cognitive stage of information processing cannot be understood unless the logically antecedent stage of environmental sampling is understood in the first place. Figure 1.1 illustrates this fundamental notion. The left box at the middle level reflects the basic assumption that the distal constructs that constitute the focus of judgment – such as health risks, student ability, a defendant's guilt, or the profitability of an investment – are not amenable to direct perception. We do not have sense organs to literally perceive risk, ability, or guilt. We only have access to samples of proximal cues (in the middle box) that are more directly assessable and that allow us to make inferences about the distal entities, to which they are statistically related. Samples of accident rates or expert advice serve to infer risk; students' responses to knowledge questions allow teachers to infer their ability in math or languages; samples of linguistic truth criteria in eyewitness protocols inform inferences of a defendant's guilt (Vrij & Mann, 2006). A nice feature of these proximal stimulus distributions is that normative rules of statistics allow us to monitor and control the process, inferring the reliability from the sample's size and internal consistency and – when the proximal data are representative of a domain – even the validity of the given stimulus information.

Regardless of how valid or reliable the environmental input is, it constrains and predetermines the subsequent cognitive judgment and decision process. The accuracy of a health expert's risk estimate, a teacher's student evaluation, and a judge's guilt assessment depend on the diagnostic value of the cue samples used to infer risk, ability, or guilt. The accuracy and confidence of their judgments and decisions depend primarily on the quality of the sampled data. Resulting distortions and biased judgments need not reflect biases in human memory or reasoning; such biases may already be inherent in the environmental sample with which the cognitive process was fed.

Indeed, the lessons taken from the entire research program of the Kahneman–Tversky tradition can be revisited and revised fundamentally from a sampling-theoretical perspective. Illusions and biases may not, or not always, reflect deficits of human memory or flawed heuristic processes within the human mind. They may rather reflect an information transition process that is anchored in the environment, prior to all cognitive operations. Samples of risk-related cues may be deceptive or lopsided; too small a sample of student responses may be highly unreliable; the defendant's sample of verbal utterances may be faked intentionally. Considered from a broader cognitive-ecological perspective, bounded rationality is not merely limited by memory restrictions or cognitive heuristics reflecting people's laziness. Judgments and decisions in the real world are restricted, and enabled, by cognitive as well as ecological limitations and capacities. For instance, risk estimations – concerning the likelihood of contracting Covid-19 or being involved in car accidents – are not just restrained by wishful thinking or ease of retrieval (Block et al., 2020; Combs & Slovic, 1979). They also depend on a rational answer to the question: What sample affords an unbiased estimate of my personal risk of a disease or accident? Should it be a sample of the entire world population, a sample of people in my subculture, or a biographical sample of my own prior behavior? As the example shows, there is no alternative to devising a heuristic algorithm for risk estimation. Heuristics are sorely needed indeed, not just for the human mind but also for machine learning, and expert and robot systems (Fiedler et al., 2021).

1.2 Historical Review of Origins and Underpinnings of Sampling Approaches

The information transition process that underlies judgments and decisions can be decomposed into two stages (see Figure 1.1): an ecological sampling stage and a cognitive processing stage. While traditional cognitive research was mainly concerned with the processing of stimulus input within the individual's mind (attention, perception, encoding, storage, retrieval, constructive inferences) of stimulus cues, the ecological input to the cognitive–decision stage reflects a logically antecedent sampling stage, which takes place in the environment. Judgment biases and decision anomalies that were traditionally explained in terms of retrieval or reasoning biases during the cognitive–decision stage may already be inherent in the stimulus input, as a consequence of biased sampling in the environment, before any cognitive operations come into play. Biased judgments and decisions can thus result from fully unbiased mental operations applied to biased sampling input. Conversely, unbiased and accurate estimates may reflect the high quality of information from certain environments.

1.2.1 Methodological and Meta-Theoretical Assets

The causal sequence (of sampling as an antecedent condition of cognitive processing) and the normative-statistical constraints imposed on the sampling stage jointly explain the beauty and fertility, and the theoretical success of sampling approaches. As in psychophysics, an analysis of the samples of observations gathered in the information search process imposes strong constraints on the judgments and decisions informed by this input. Statistical sampling theory imposes distinct normative constraints (in terms of sample size, stochastic independence, etc.) on how inferences from the sample should be made. Both sources of constraints together lead to refined hypotheses that can be tested experimentally. Because the causal and statistical constraints are strong and clear-cut, the predictions tested in such experiments are cogent and nonarbitrary, and, not by coincidence, empirical findings often support the a priori considerations. Indeed, replication and validation do not appear to constitute serious problems for sampling research (Denrell & Le Mens, 2012; Fiedler, 2008; Galesic et al., 2018).

1.2.1.1 Recording the Sampled Input

Having a measure of the sampling input in addition to the judgments and decision in the ultimate dependent measure offers a natural candidate for a mediational account of cognitive inferences relying on the sample. Comparing the recorded sample to the ultimate cognitive measure provides a way to disentangle the two processes. A causal origin of a judgment or decision effect that is already visible in the actuarial sample must originate in the environment, before cognition comes into play. Evidence for a genuine cognitive influence (e.g., selective retrieval or an anchoring bias) requires demonstrating a tendency in the cognitive process that is not yet visible in the recorded sample.

Let us illustrate the methodological advantage of having a record of the sample with reference to recent research on sample-based impression judgments. Prager et al. (2018) had participants provide integrative likeability judgments of target persons described by samples of $n = 2$, 4, or 8 traits drawn at random from a universe defined by an experimentally controlled distribution of positive and negative traits. Each participant provided 36 impression judgments, nine based on random samples drawn from each of four universes of extremely positive, moderately positive, moderately negative, and extremely negative sets of traits, selected in careful pilot testing. Across all participants and trials, impression judgments were highly predictable from the recorded samples of traits. Not only the positive versus negative valence and extremity of the universe from which the stimulus traits were drawn, but also the deviations of the random samples from the respective universe strongly predicted the ultimate impression judgments. Consistent with Bayesian updating principles, impression extremity increased with increasing *n*. Altogether, these findings provided strong and regular support for the (actuarial) stimulus sample as major determinant of person impression (Asch, 1946; Norton et al., 2007; Ullrich et al., 2013).

However, in spite of their close fit to the sampled input, the impression judgments were also highly sensitive to the structure of the environment, specifically, the diagnosticity of the information. The diagnosticity of a trait is determined by the covariation of features in the environment and can be defined in the same way as in a likelihood ratio in Bayesian updating; a trait is diagnostic for a hypothetical impression (e.g., for the hypothesis: likable person) to the extent that it is more likely to occur in a likable than a nonlikable person.[1]

Holding the valence scale value of the sampled traits constant, diagnostic traits exerted a stronger influence on person judgments than nondiagnostic traits. Diagnosticity was enhanced if a trait was negative rather than positive (Rothbart & Park, 1986); if a trait referred to negative morality or positive ability rather than positive morality or negative ability (Fiske et al., 2007; Reeder & Brewer, 1979); if a trait was infrequent rather than

[1] Thus, in Bayesian notation, a trait is diagnostic to the extent that the likelihood ratio LR = $p(\text{trait} \mid H_{\text{likeable}}) / p(\text{trait} \mid H_{\text{not-likeable}})$ exceeds 1.

frequent (Prager & Fiedler, 2021); or if a trait's distance from other traits in a semantic network was high (Unkelbach et al., 2008). However diagnosticity was operationalized, the resulting impression of a target person was not fully determined by the average valence scale value of the traits recorded in a sample but depended on the diagnosticity of the sampled traits. Adding a diagnostic trait had a stronger impact on a growing impression than adding a nondiagnostic trait of the same valence.

Further evidence of how people actively interpret the observed samples will be provided later. Suffice it here to point out the advantage of a research design with a twofold measure for the sampling input on one hand and for the cognitive process output on the other hand. Let us now turn to the second major asset of the sampling-theory approach, namely, the existence of normative constraints imposed by statistical sampling theory on the information transition process. To the extent that judgments and decisions are sensitive to such distinct normative constraints, which often exceed intuition and common sense, this would provide cogent evidence for the explanatory value of sampling theories.

1.2.1.2 Impact of Sampling Constraints

The keywords in the lower left of Figure 1.1 refer to a number of subtle sampling constraints, which are firmly built into the probabilistic environment. For instance, in a world in which many frequency distributions are inherently skewed, probability theory constrains the probability that a sample reveals a dominant trend, for instance, that a sample reflects the relative frequency of lexical stimuli, animals, or causes of death. Skewed distributions are highly indicative of moral and material value. Rare objects tend to be more precious than common things (Pleskac & Hertwig, 2014); scarcity increases the price of economic goods. Abnormal or norm-deviant behaviors are less frequent than normal or norm-abiding behaviors. Likewise, skewness is indicative of psychological distance. Frequently encountered stimuli more likely belong to temporally, spatially, socially, close and probable origins than infrequent stimuli, which are indicative of distant origins (Bhatia & Walasek, 2016; Fiedler et al., 2015; Trope & Liberman, 2010). In any case, normal variation in distance, density, resolution level, and perspective can open up a variety of environmental information.

Small samples from skewed distributions are often unrepresentative of the underlying distribution and this can lead to seemingly biased judgments. Suppose, for example, that the population probability of a success is 0.9. In a small sample of five trials, an agent will most often observe a proportion larger than 0.9; the probability of observing five successes in five trials is 0.6. It is 0.77 if the probability of a success is 0.95. Thus, if judgments are sensitive to experienced proportions, most agents will overestimate the success probability. To be sure, agents may be more sophisticated and understand that small samples can be unrepresentative. Suppose an agent believes that all probabilities between zero and one are equally likely (a uniform prior distribution) and uses this information in combination with the observed proportion. Such a Bayesian agent will estimate the true success probability to be lower than the population proportion of 0.9. Having observed five successes in five trials, this Bayesian agent will estimate the success probability to be only 0.86.[2] Thus, normative-statistical laws not only justify that sample proportions can deviate from true probabilities in the population but also specify predictions of how sample-dependent estimations can be expected to deviate from population parameters.

It is no wonder then that decisions about risk-taking differ substantially between settings where the winning probability of a lottery is described numerically versus when a sample of outcomes is experienced extensionally – the so-called description–experience gap (Hertwig et al., 2004). Statistical sampling theory as an integral part of a cognitive-ecological approach can therefore offer a viable explanation of many findings related to the description–experience gap (Fox & Hadar, 2006; Rakov et al., 2008).

The importance of skewed sampling distributions also inspired a prominent finding by Kareev (2000). Assuming an actually existing (population) correlation of, say, $\rho = .50$, the majority of observed correlations r in restricted samples from this population is higher than ρ. (Undoing this asymmetry of the sampling distribution of r-statistics is the purpose of the common Fisher-z transformation). Kareev (2000) showed that the tendency of r to exaggerate existing correlations reaches a maximum at $n = 7 \pm 2$, suggesting that the evolution may have prepared *Homo sapiens* with a memory span that maximally facilitates the extraction of existing regularities. Regardless of the viability of Kareev's vision (see Juslin & Olsson, 2005, for a critical note), it clearly highlights the fascinating ability of sampling theories to inform creative theorizing in cognitive-ecological context.

2 According to the Bayesian rule of succession (Costello & Watts 2019); the underlying probability of the dominant outcome is $p = (n_{\text{dominant}} + 1)/(n_{\text{total}} + 2)$. Thus, observing $n_{\text{dominant}} = 8$ dominant outcomes in a sample of $n_{\text{total}} = 10$ implies $p = .75$. Observing the same proportion $n_{\text{dominant}} = 4$ in a sample of $n_{\text{total}} = 5$ implies $p = .71$.

Let us now move from unsystematic sampling error (around *p* or ρ) derived from statistical sampling theory to systematic sampling biases lying outside the domain of statistics.[3] Some behavioral laws are so obvious and universal that one hardly recognizes their statistical consequences. For example, Thorndike's (1927) law of effect states that responses leading to pleasant outcomes are more likely repeated than responses leading to unpleasant outcomes. In other words, organisms are inclined to sample more from pleasant than from unpleasant sources. A hot-stove effect motivates organisms to stop sampling from highly unpleasant sources (e.g., a restaurant where one got sick). Such a simple and self-evident preference toward hedonically positive stimuli was sufficient to inspire a series of highly influential simulations and experiments that opened up completely novel perspectives on behavior regulation (Denrell, 2005; Denrell & Le Mens, 2007, 2011; Fazio et al., 2004). The tendency to stop sampling from negative targets and more likely continue sampling from positive targets implies that negative first impressions are less likely corrected than positive first impressions. Long-term negativity biases may be the result of such a simple and incontestable hedonic bias.

An example of another effect that has been prematurely taken for a cognitive bias refers to the seminal work on heuristics and biases by Tversky and Kahneman (1973). In their famous availability heuristic, they postulate a cognitive bias to overestimate the frequency or probability of easily retrievable events. Thus, a bias in frequentist judgments is attributed to a cognitive bias to overrate information that easily comes to one's mind. Hardly anybody ever contested that the bias may be already apparent in the *sampling stage*, well before a retrieval bias may come into play, although this possibility was discussed from the beginning. For instance, the erroneous tendency to rate murder more frequent than suicide, to overrate lightning and to underrate coronary disease as causes of death, need not reflect a retrieval bias but a bias in newspaper coverage (Combs & Slovic, 1979). Thus, prior to cognitive retrieval processes, newspapers or the information environment are more likely to report on murder than on suicide, on lightning than coronary disease, and this preexisting sampling bias may account for availability effects. Even when every cause of death reported in the media is equally likely to be retrieved, biased media coverage may well account for biased probability estimates. A critical examination of the literature reveals, indeed, that countless experiments

[3] Note however, that statistical rules are essential to distinguish systematic biases from unsystematic (merely stochastic) error.

on the availability heuristic have provided little evidence for memory retrieval proper.

1.2.2 Properties of Proximal Samples

So far, we have seen that merely analyzing the statistical properties or the hedonic appeal of the environment opens up alternative explanations of various psychological phenomena as well as genuine innovations that had been never discovered without the sampling perspective. The following discussion of the properties of the proximal samples implanted by the distal world leads to further insights about the beauty, fertility, and the explanatory power of the sampling approach.

One important and common property of observations based on noisy samples is regression to the mean. If the observed value, X' is higher than Mean_X, then the true value X is likely lower than X', but if the observed value, X' is lower than Mean_X, then the true value X is likely higher than X' (see Figure 1.2; for precise definitions see Samuels [1991], and Schmittlein [1989]). This property holds for many distributions and implies that observed values diverge regularly and in predictable ways from true values.[4] Regressiveness increases with the amount of noise, or error variance. To illustrate this, consider two normally distributed random variables, X' and X, where X' is a noisy observation of X (i.e., $X' = X + e$, where e is an error term). Suppose, for simplicity, that the variables are standardized to z scores with zero mean and variance: $z_X = (X - \mathrm{Mean}_X)/\mathrm{SD}_X$. Then the expected z_X given an observed standardized $z_{X'}$ value is $E[z_X|z_{X'}] = r_{X,X'}\, z_{X'}$. Whenever the correlation between X' and X is less than 1, the best estimate of z_X given $z_{X'}$ is less extreme than $z_{X'}$. Specifically, if the correlation is $r_{X,X'} = 0.50$, the expected population values are only half as extreme as the observed values; if $r_{X,X'} = 0.75$, the expected population values shrink by one-fourth to 75 percent of the observed deviation from the mean. Thus, observed values diverge from expected values in predictable ways.

While regression is not a bias but a reflection of noise in the probabilistic world, it can create what appears to be a bias. An agent that reports the raw observed sample value as their estimates – like ignoring effects implied by regression to the mean – will make systematically too extreme judgments, as compared to the long-run expected value. Alternatively, an agent may take expectable regression effects into account and make

[4] Regression to the mean does not hold for all distributions, however. There may be regression to the median or sometimes even regression to the extremes, see Schmittlein (1989).