

## Memory, metamemory, and conditional statistics

Robert A. Bjork and Thomas D. Wickens

Department of Psychology, University of California, Los Angeles, CA  
90095-1563. rabjork@psych.ucla.edu; twickens@psych.ucla.edu

**Abstract:** Koriat & Goldsmith's distinction between encoding processes and metamnemonic decision processes is theoretically and practically important, as is their methodology for separating the two. However, their accuracy measure is a conditional statistic, subject to the unfathomable selection effects that have hindered analogous measures in the past. We also find their arguments concerning basic and applied research mostly beside the point.

Koriat & Goldsmith (K&C) provide an interesting characterization of alternative approaches to the study of memory. They remind us of some important matters concerning how to look at our data in either laboratory or real-world contexts, and they explicate nicely the importance of metamnemonic decision processes as contributors to observed performance. However, we find their characterization of basic and applied research to be disputable and largely irrelevant to their broader issue, and we want to comment on their distinction between free and forced measures of memory and on the limits to their conditional measure of memory accuracy.

In the free-report experiments K&C describe (conventional free recall the most familiar example), subjects determine what they report of the information they recall or think they may recall. In such experiments, information can be reported correctly or incorrectly, or omitted (RC, RI, and O, respectively). With three categories, performance cannot be completely captured by any single index. Quantity of retention, in the sense used by K&C, is measured by  $P(RC)$ , and conditional accuracy by  $P(RC|R) = P(RC)/P(RC \cup RI)$ . Neither measure captures how much is reported and how much omitted, or, more importantly, the content of the omitted material. We want to amplify the real-world and theoretical implications of this point.

K&C imply that omissions are not a serious problem in real-world contexts like witness memory. In section 4.2.2.2, for example, they mention a hypothetical witness who correctly recalls that individuals A, B, and C were present at a crime scene but fails to mention individuals D and E. From testifying in a number of cases, we can say with confidence that if other evidence establishes that D and E were at the scene, the witness will not be considered "accurate" by the jury. The failure of the witness to remember salient aspects of the criminal episode leads juries to lose confidence in what the witness does report.

From both a practical and a theoretical standpoint, it is essential to recognize that K&C's outward-bound accuracy measure is a conditional statistic. It is subject, therefore, to the complex and confusing selection artifacts that have always bedeviled such measures. If different material is reported in different experimental conditions, then such conditional statistics cease to be comparable. For example, in experiments where study conditions (e.g., level of processing) or characteristics of materials (e.g., frequency, concreteness, meaningfulness) are manipulated, each of which has large and known effects on later memory performance, looking at accuracy alone and ignoring what was actually reported can lead to some very strange conclusions. Specifically, subject-by-item selection effects can make the conditional measure  $P(RC|R)$  higher for poor study conditions, more difficult materials, and less-alert subjects.

In the orthographic condition of a levels-of-processing experiment, for example, judging whether the word *pumpkin* is in uppercase letters normally produces abysmal recall of that word, but if *pumpkin* is your special name for your girlfriend, you may be unable to avoid the type of self-referential processing known to produce excellent recall. If you recall only a few such items, then you will show good outward-bound accuracy, but we would not want to conclude that your memory is "accurate" overall. Similarly, a witness who had a good opportunity to view a complex episode and reports many things about it may have lower output-bound accuracy than a witness who had a poor opportunity to observe and reports only a few salient facts. The latter witness should not necessarily be perceived as more credible.

From this perspective, a forced-report procedure is one way to eliminate the selection issue and the problems of conditional fixation. This simplification is one of the reasons these procedures are widely adopted, although we do not argue that it makes them superior to free report. In this regard, it is interesting to note that courtroom testimony is seldom free report. Witnesses are rarely permitted to give narrative answers and are often forced to answer questions. As the American legal system has evolved, it is deemed important that both the prosecution and the defense get answers to everything they ask, yielding something close to a quantity measure.

We also take issue with K&C's version of the storehouse metaphor. For expository purposes, they describe a simplified version of this metaphor, but their version distorts its current form, which we see as closer to a library metaphor. An overworked librarian (some kind of central executive in memory) has a complex and individualistic filing system based on experience. Additions to the library are filed in terms of where they fit, as determined by their relationships to what is already shelved. If the librarian is too taxed by other demands when new materials are received, or if the new materials defy ready classification in terms of existing materials, then the new materials do not get shelved in permanent storage at all. At the time books are requested, the librarian uses the cues provided (title, topic, author, publisher, etc.) to try to find a given volume, but those cues may underspecify the target volume or may define a competing alternative better than the desired volume. The organization of the library changes over time, induced by the continuing addition and retrieval of materials, so that volumes that are retrievable at one time may not be retrievable at another. This version of the "storehouse" metaphor captures more of the dynamic and content-specific character of memory than does the less realistic version cited by K&C.

Finally, we find the laboratory-real-world aspect of the metaphor somewhat irrelevant. To make the enormous advances they did, physicists, chemists, and biologists worked with simple materials - smooth balls and inclined planes, not jagged rocks falling down rugged hillsides. Why then should we study memory only in the full social, emotional, and semantic complexity of everyday life? Real-world situations suggest general principles that are worked out in simplified domains. Controlled research on memory with simple materials has not misled us about basic functional relationships. The phenomena and principles generated by laboratory research - such as the effects of spacing, the number and nature of study trials, the length and content of retention intervals, the differences between recall and recognition or between concrete and abstract materials, the effects of meaningfulness, the differences between discrimination and criterion ( $d'$  and  $\beta$ ), and the types of interference and transfer - have routinely been replicated in real-world memory settings, often in enlarged form.