Is Science an Evolutionay Process? Evidence from Miscitation of the Scientific Literature

Kim J. Vicente

Perspectives on Science, Volume 8, Number 1, Spring 2000, pp. 53-69
(Article)

Published by The MIT Press

Is Science an Evolutionary Process? Evidence from Miscitations of the Scientific Literature

Kim J. Vicente
University of Toronto

This article describes a psychological test of Hull's (1988) theory of science as an evolutionary process by seeing if it can account for how scientists sometimes remember and cite the scientific literature. The conceptual adequacy of Hull's theory was evaluated by comparing it to Bartlett's (1932) seminal theory of human remembering. Bartlett found that remembering is an active, reconstructive process driven by a schema that biases recall in the direction of prototypicality and personal involvement. This account supports Hull's theory of science because it shows that the characteristics of reconstructive remembering are consistent with the generic properties of an evolutionary process. The empirical adequacy of Hull's theory was evaluated by comparing the predictions made from this evolutionary viewpoint against evidence from the history of science. Six cases studies of well-known psychological experiments that had been subject to repeated miscitation errors were collected and reviewed. All six case studies revealed a systematic pattern of distortions that is consistent with the schema-induced biases of reconstructive remembering. These findings support Hull's claim that science is an evolutionary process with scientists as interactors, scientific beliefs as replicators, and schemata as means for that replication.

Introduction

Scientific studies of science are few and far between (Hull, 1998). One reason is that it is difficult to study scientific activity in a representative fash-

The writing of this paper was sponsored by grants from the Natural Sciences and Engineering Research Council of Canada. I would like to thank Bill Brewer and David Hull for their constructive criticisms and help, and Reid Hastie for bringing some of the case studies to my attention. Correspondence and requests for reprints should be addressed to Dr. Kim J. Vicente, Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, Ontario, Canada M5S 3G8 (E-mail: benfica@mie.utoronto.ca, URL: www.mie.utoronto.ca/labs/cel/).

Perspectives on Science 2000, vol. 8, no. 1

ion using the traditional methods of science, such as highly controlled laboratory experiments—scientists are not randomly assigned to their careers. Although it may be difficult to collect "hard" empirical evidence, it is possible to gather descriptive evidence that can constrain viable theories of scientific activity. As this article shows, case studies from the history of science—psychology, to be more precise—can serve precisely this function (*cf.* Simonton 1997).

Another reason for the comparatively few empirical investigations of scientific activity is the rarity with which philosophers of science have proposed theories of science that are empirically testable. Hull's (1988) evolutionary account of the social and conceptual development of science is one such theory. In this article, predictions made from that theory are compared with empirical evidence about the professional activity of research scientists.

Science as an Evolutionary Process

Hull's (1988) thesis is that science is an evolutionary process, much like biological evolution, except with different entities. The correspondence between these two seemingly different processes can be made clear by describing evolution in a generic fashion, without making specific reference to the particular case of biology. Hull proposed the following four concepts to achieve this goal.

1. *Replicator:* "an entity that passes on its structure largely intact in successive replications" (p. 408).
2. *Interactor:* "an entity that interacts as a cohesive whole with its environment in such a way that this interaction *causes* replication to be differential" (p. 408, emphasis in original).
3. *Selection:* "a process in which the differential extinction and proliferation of interactors cause the differential perpetuation of the relevant replicators" (p. 409).
4. *Lineage:* "an entity that persists indefinitely through time either in the same or an altered state as a result of replication" (p. 409).

By instantiating these generic concepts in the particular context of biology, an account of biological evolution is obtained: genes can be replicators; organisms can be interactors; natural selection can be a form of selection; and species can be lineages. Other entities can serve each of these roles (e.g., genes can also be interactors), but these examples suffice to illustrate the gist of the concepts.

Hull's (1988) four generic concepts can also be instantiated in an analogous fashion for scientific activity: scientific knowledge and beliefs can be replicators; scientists can be interactors; competition between scientists

can be a form of selection; and scientific paradigms can be lineages. Hull puts forth a detailed body of evidence in support of his theory, most of it coming from the social interaction between scientists.

However, the view that science is an evolutionary process can also be tested from a psychological perspective. There are many ways in which this insight can be explored (e.g., Hull 1988, pp. 368–371). Perhaps the most obvious way is to investigate the psychology of scientific invention. Although there is merit to this approach, it is a difficult path to pursue empirically because psychological evidence about invention processes is difficult to obtain. Fortunately, science evolves, not just through the rare quantum leaps of the truly creative, but also by the frequent mundane dissemination of ideas among scientists. This article is focused on the memory of scientists, a phenomenon that is easier to observe than scientific invention. In the normal course of teaching and conducting research, scientists frequently rely on their memory of the scientific literature to pass on their knowledge and beliefs to both students and colleagues. If science is indeed an evolutionary process, as Hull claims, then we would expect that this process of disseminating scientific knowledge via the memory of scientists should exhibit properties that are consistent with those of a generic evolutionary process, as described above. Fortunately, this hypothesis is empirically falsifiable. Before it can be put to the test, however, we need to know more about the characteristics of human memory.

Memory or Remembering? Noun *vs.* Verb Theories

Traditionally, human memory has been conceptualized using a storehouse metaphor. Bluntly, memory is a "place" in your brain where you "encode" traces, "store" them for later use, and then "retrieve" them when they are required. This view of memory was bolstered by the computer metaphor of mind that became prominent in psychology beginning in the 1950s and 1960s. Encoding, storage, and retrieval functions in computer memory provided a model for human memory (e.g., Atkinson & Shiffrin 1968; Waugh & Norman 1965). One can still find such theories being proposed in *Psychological Review,* the most prestigious of all psychological journals (e.g., Hintzman 1986; Logan 1988).

At the risk of greatly oversimplifying matters, such theories of human memory can be considered as "noun theories" in the sense that they view memory as a thing or a place. This view of memory poses problems for Hull's (1988) evolutionary theory of science. After all, noun theories of memory are quite different from evolutionary phenomena. Species or organisms are not "stored" in genes, and conversely, replication does not play any meaningful role in the storehouse view. Does this mean that Hull's theory is falsified by psychological data? Or is there some other

view of human memory that is more compatible with an evolutionary perspective?

Interestingly, an alternative view can be found as far back as 1932, the year that Bartlett (1932) published his well-known treatise on reconstructive remembering. Note the change in terminology; "memory" has been replaced by "remembering" (the title of Bartlett's book). This is not an accident since Bartlett viewed remembering as an active *process,* rather than as a thing or a place. In short, his was a "verb" theory, not a "noun" theory (*cf.* Kolers & Roediger 1984; Morris, Bransford, & Franks 1977).[1]

An intuitive appreciation for Bartlett's (1932) theory can be illustrated by describing one of his best-known series of experiments. In those studies, a story was given to a participant to remember, and then that participant's recollection was passed along in writing for someone else to remember, and so on. As shown by the example in table 1, the story became increasingly distorted as it was serially disseminated to more and more people. The key point, however, is that the distortions were not random or chaotic, but instead occurred in a regular pattern. The following features in participants' recall were observed: condensation, elaboration, invention, simplification, integration towards greater coherence, and omission of qualifications. Globally, these patterns tended to make the original story simpler and more meaningful to participants, thereby making it easier to remember. Similar findings were also obtained by Allport and Postman (1945) in their investigation of the psychology of rumor.

What type of psychological mechanism could account for these findings? According to Bartlett (1932), these changes in the text were generated by an active process of *reconstructive remembering* wherein global knowledge structures (schemata) interact with incoming instances to produce systematic distortions in recall (*cf.* Brewer & Nakamura 1984). That is, people retain a general impression of the original event and then use that impression to generate a simple, yet plausible, account of the forgotten details. In Bartlett's (1932) own words: "Remembering is not the re-excitation of innumerable fixed, lifeless and fragmentary traces. It is an imaginative reconstruction, or construction, built out of the relation of our attitude towards a whole active mass of organized past reactions or experience" (p. 213); "So long as the details which can be built up around

1. The astute reader may have noticed that Hull's (1988) theory was originally formulated in terms of nouns (e.g., replicator, interactor, selection, lineage), not verbs. Interestingly, Hull (personal communication, October 1997) has reformulated his theory in terms of verbs in order to better capture its process-oriented nature. While this development has occurred independently of the ideas presented in this article, it strengthens and emphasizes the connection between science as an evolutionary process and reconstructive remembering.

Table 1. An example of the results from one of Bartlett's (1932, pp. 166–167) famous experiments on remembering. The first passage is the original one. The second passage is the first participants' written recall of the original. This recall then served as the passage to be remembered by the second participant, and so on. The third passage shows the written recall of the eleventh participant in the sequence.

Original Passage:

(i) Modification of Species

One objection to the views of those who, like Mr Gulick, believe isolation itself to be a cause of modification of species deserves attention, namely, the entire absence of change where, if this were a *vera causa,* we should expect to find it. In Ireland we have an excellent test case, for we know that it has been separated from Britain since the end of the glacial epoch, certainly many thousand years. Yet hardly one of its mammals, reptiles or land molluscs, has undergone the slightest change, even though there is certainly a distinct difference of environment, both inorganic and organic. That changes have not occurred through natural selection is perhaps due to the less severe struggle for existence owing to the smaller number of competing species; but if isolation itself were an efficient cause, acting continuously and cumulatively, it is incredible that a decided change should not have been produced in thousands of years. That no such change has occurred in this and many other cases of isolation seems to prove that it is not itself a cause of modification.

Reproduction of First Participant:

The Modification of Species

The objection put forward by Mr Gulick, that isolation is not a sufficient cause for the modification of species, deserves attention. For there are certain things which, if this were a *vera causa,* are not as we should have expected them to be. In Ireland we have an excellent test-case. We find there mammals, reptiles and molluscs like those in this country, yet it was separated from Britain during the glacial period for many thousands of years.

Even if modification of species were a result of isolation . . .

Reproduction of Eleventh Participant:

Mr Garlick says that isolation is the result of modification. This is the reason that snakes and reptiles are not found in Ireland.

[the remaining impression] are such that they would give it a 'reasonable' setting, most of us are fairly content, and are apt to think that what we build we have literally retained" (p. 176). The active, reconstructive nature of everyday memory has since been well documented (e.g., the various studies in Neisser 1982, and Neisser & Winograd 1988).

Note that this "verb" theory of remembering is quite distinct from the "noun" theory of memory represented by the storehouse metaphor. With noun theories, the stimulus to be recalled, the trace that is said to be stored in memory, and the output of recall are all of the some qualitative type. From this perspective, these entities are merely shuffled around from one "place" to another. With verb theories, however, there is a qualitative difference between the stimulus material to be remembered and the schema that serves as the psychological mechanism for remembering. The schema is not a trace, copy, or record of the stimulus, but is instead the means by which the recall is actively constructed. Thus, rather than retrieving a trace from memory, remembering is viewed as a generative process by which an entity of a new sort is literally constructed on the basis of a schema. As a result, the schema and the product of remembering are of qualitatively different types. The latter cannot be retrieved from the former; it must be built through an active process.[2]

Bartlett (1932) himself pointed out the contrast between the schemata and storehouse views very clearly: "A storehouse is a place where things are put in the hope that they may be found again when they are wanted exactly as they were when first stored away. The schemata are . . . living, constantly developing, affected by every bit of incoming sensational experience of a given kind. The storehouse notion is as far removed from this as it well could be" (p. 200). In other words, remembering is an active process, not of retrieving what was encoded and stored, but of reconstructing details on the basis of a schema.

For the purposes of this article, the important implication of the contrast between noun and verb theories is that generative schema processes have a much stronger relationship to a generic evolutionary process

2. It is important to note that this characterization applies to Bartlett's (1932) original conception of schema, but not to most recent incarnations of the same concept. As Caramelli (1987) has pointed out, contemporary cognitive science and artificial intelligence conceptions of "schema" are quite different from Bartlett's and, ironically, constitute a distortion of his original ideas. As the quotes throughout this paper indicate, Bartlett seemed to have had in mind a dynamic, fluid process of self-organization that is perhaps better captured by modern dynamical systems theory models (e.g., Thelen & Smith 1994; Kelso 1995) rather than by the static structures with slots and variables proposed by cognitive science and artificial intelligence researchers.

than do the elements in the storehouse view of memory (*cf.* Plotkin 1994). Schemata play a role that is functionally similar to that of DNA. Just as DNA does not store copies of organisms, schemata do not store traces of what is remembered. Rather, schemata are the means by which ideas and knowledge are replicated (i.e., actively reconstructed), just as DNA is the means by which genes are replicated. Again, the means and the output of the evolutionary process are of a qualitatively different type because the latter is, not retrieved, but generated (or built) from the former.

A major difference, of course, is that the replication reliability rate of DNA is far superior to that of schemata. As Bartlett (1932) observed, remembering is subject to a number of biases that can frequently lead to systematic distortions in recall. These distortions (i.e., errors in replication) are somewhat predictable, if we know some general characteristics of the schema that people are using for recall (i.e., the means for replication) and the people (i.e., the interactors) themselves. This knowledge allows us to make predictions that can be empirically tested, thereby providing a means for evaluating Hull's (1988) theory of science as an evolutionary process from a psychological perspective.

As mentioned earlier, the psychological phenomenon addressed in this article is recall of the scientific literature by scientists. In this particular case, there are two potential sources of distortion errors. First, reconstruction is frequently biased by constraints of prototypicality. In the present case, it is hypothesized that scientists have a schema for the canonical scientific experiment which they use as a basis for reconstructive remembering. This idealized schema incorporates the belief that the perfect scientific experiment provides a clean test of a particular theory with appropriate controls and clear, unambiguous results. In this schema, the theory explains the findings, the controls reduce alternate interpretations, and the results provide clear cut evidence in support of the theory. It is further postulated that when scientists are attempting to describe a particular experiment to make a certain point or as an example for teaching purposes (e.g., in lectures or introductory texts), there will be a strong tendency for the idealized experiment schema to override the actual details of a given experiment. Second, as Bartlett (1932) pointed out, the schema that a particular person adopts for reconstruction is also biased by people's motives. In Hull's (1988) terms, replication is affected by characteristics of the interactors. In Bartlett's own words: "the mechanism of adult human memory demands an organisation of 'schemata' depending upon on an interplay of appetites, instincts, interests and ideals peculiar to any given subject" (p. 213); "The recall is then a construction, made largely on

the basis of this attitude, and its general effect is that of a *justification of the attitude*" (p. 207, emphasis added). In the cases of scientists citing the literature, it is hypothesized that this personal involvement arises because scientists frequently wish to prove a particular point. This motivation or attitude may involuntarily cause scientists to distort the results of the literature they are citing in such a way that it provides a more clear or convincing justification for the point they wish to make. Consequently, it is predicted that scientist's recall of the literature should sometimes exhibit the characteristics of schema-based reconstructive remembering. If so, then this result would support Hull's theory, since reconstructive remembering is itself an evolutionary process.

In the next section, six cases studies will be reviewed, each consisting of a well-known psychological experiment that has been repeatedly miscited in the literature. The goal is to determine if these case studies provide any empirical evidence for the predictions just made. Can reconstructive remembering, and thus Hull's (1988) theory, account for the symptoms exhibited by a sample of miscitations of the scientific literature?

Case Studies

de Groot (1946/1965)

As part of his doctoral dissertation de Groot (1946, 1965) conducted a seminal investigation of expert problem solving in chess. This work has become widely cited in the cognitive psychology and cognitive science literatures on expertise, particularly since it was brought to the attention of the international research community by Chase and Simon (1973*a,* 1973*b*). However, as Vicente and de Groot (1990) have pointed out, de Groot's original work has been subject to persistent distortions in both the primary and secondary literatures.

De Groot's (1946, 1965, pp. 321–334) best-known experiment tested chess players' abilities to reconstruct a board position. There were four participants with different levels of chess expertise (Grandmaster, Master [de Groot himself], Expert, and Class level). Participants were presented with mid- and end-game positions from actual games. The exposure times ranged from 2 to 15 seconds. De Groot found that the two more experienced players were able to score over 90% (as determined by an arbitrary scoring criterion). The recall performance of the less skilled participants was considerably worse. Years later, de Groot's result was replicated by Jongman (1968), a Ph.D. student of his, this time with exposure times of 5 seconds. Jongman and a co-worker, Lemmens, also conducted a control experiment where the pieces were randomly placed on the board. The

memory advantage found on authentic positions disappeared in this random condition. Unfortunately, this study was never published. In 1973, Chase and Simon (1973*a,* 1973*b*) replicated the original results of de Groot with novice and Master level participants and exposure times of 5 seconds. They also replicated the unpublished random condition findings of Jongman and Lemmens. This was the first published memory recall study that incorporated a random control condition.

Beginning with the two papers by Chase and Simon (1973*a,* 1973*b*), de Groot's (1946, 1965) memory experiment has been repeatedly miscited (see Vicente & de Groot 1990 for a detailed account). The most frequent error is one of attribution; many authors incorrectly state that de Groot conducted the random control condition. The postulated schema for canonical experiments explains why scientists consistently make this error. The random control condition is essential to the interpretation of the findings. After all, if the random control is not included, then experts could simply have a better overall memory than lesser players. Thus, scientists remember the gist of the chess experiment and then use that gist to reconstruct the details. Since the gist is not meaningful without the random control, scientists erroneously recall that de Groot must have conducted this control in addition to the condition with meaningful chess positions. In short, scientists report the experiment as it should have been conducted rather than how it really was conducted.

Most of the other patterns in the citation errors of de Groot's (1946, 1965) work are also consistent with the schema-based reconstruction hypothesis. For example, several authors have incorrectly reported that there were two groups of participants in the experiment, novices and experts. The vast majority of contemporary expertise studies involve a comparison between extreme groups, so the novice participant group provides a canonical control group for studies of expertise. Also worthy of note is that the gist of the research (i.e., that chess experts remember more than weaker players when the stimulus is meaningful, but not when the stimulus is random) is always recalled accurately. Since this gist is the core of the postulated canonical schema, we would expect it to be retained. The only recurring citation error that does not seem to be caused by reconstructive remembering is the tendency to report the exposure time as being 5 seconds. This error is likely a source confusion error (*cf.* Schacter, Harbluk, & McLachlan 1984), since Chase & Simon's (1973*a,* 1973*b*) exposure times were 5 seconds.

In summary, the pattern of miscitations of de Groot's well-known work is generally consistent with the "signature" of schema-based reconstructive remembering.

Watson and Rayner (1920)

Watson and Rayner's (1920) behaviorist study of Little Albert is perhaps the best known example of misrepresentation in the psychological literature (*cf.* Cornwell & Hobbs 1976 ; Harris 1979; Prytula, Oster, & Davis 1976). Like the de Groot study and the other experiments to be described below, the case of Little Albert is well known to psychologists. In fact, Harris (1979) reports that the study is one of the most widely cited experiments in psychology textbooks. The question addressed by Watson and Rayner's (1920) study was whether it was possible to condition various types of emotional response. They carried out their experiment with an infant, Albert B.

The numerous and varied misrepresentations of this experiment have been well documented, most notably by Harris (1979). The errors include misrepresentations of both the methods and the results. In trying to come up with an explanation for these miscitations, Harris (1979) observed that some of the errors seem to arise from the fact that authors have a tendency to make experimental evidence conform with their favorite theories. In support of this claim, Harris points out that two groups of scientists, one advocating behavior therapy and the other advocating preparedness theory, each misrepresented the Little Albert experiment but in different ways. Although the errors committed by the two groups were different, they could be explained in a systematic fashion. More specifically, the citation errors committed by each group of scientists had the effect of making their distorted accounts more in line with their respective theoretical suppositions. In the early 1960s, behavior therapists incorrectly claimed that Little Albert developed a rat phobia. Ten years later, proponents of preparedness theory incorrectly stated, among other things, that Watson himself did not become an aversive stimulus to Albert. This pattern of findings can be interpreted in terms of schema-based reconstructive remembering, since the information that was recalled from the experiment seems to have been distorted by the personal involvement associated with the differing theoretical views of the respective scientists.

In addition, some authors have stated that Little Albert was desensitized after the experiment, a standard procedure in experiments of this sort, when in fact he was not. This error provides an example of authors adopting an idealized experiment schema to reconstruct the details of an original experiment that did not conform to the idealized schema. Scientists describe the experiment the way it should have been done, not the way it was actually done. Note that this distortion cannot be accounted for by biases resulting from an individual's personal involvement or specific theory. Rather, it is related to the general way in which experiments of

this type tend to be conducted. This observation provides further support for the schema-based reconstruction hypothesis.

Allport and Postman (1945)

Allport and Postman (1945) conducted a classic study investigating the psychology of rumor that was very similar to Bartlett's (1932) experiments with the method of serial reproduction. In the Allport and Postman experiments, a slide was projected onto a screen while six or seven participants, who had not seen the picture, waited in a room nearby. The first participant entered and took a position where he or she could not see the screen. Someone in the audience then described the picture to the participant. A second participant entered the room and stood beside the first participant who proceeded to tell the second participant everything that the original participant could remember about what he or she had been told was in the picture. A third participant entered to hear the story from the second participant, and the procedure was repeated several times. The experimenters found that the participants' reports often became increasingly distorted as the message was passed along to each successive participant.

It is indeed ironic that this study of the psychology of rumor has also been subject to misrepresentation. Treadway and McCloskey (1987) have pointed out that both the methods and the results of the Allport and Postman study have been repeatedly distorted in the psychological literature. Furthermore, several types of citation errors can also be found in the legal literature on eyewitness testimony, apparently because legal scholars relied on the erroneous secondary citations in the psychological literature. The most blatant error, committed in the psychological and legal literature and in the courtroom as well, was the claim that the participants who were responsible for the distortions had viewed the picture. In fact, the participants' reports were based on what they had heard from others, not on what they had seen.

These distortions of the Allport and Postman study seem to be another example of schema-based reconstructive remembering. The experiment, as described in both legal and psychological secondary sources, seems to have been distorted by an idealized experiment schema based on the theoretical belief that expectations can affect direct visual perception. The actual Allport and Postman study cannot support such a claim because the participants in that study were not eyewitnesses. Once again, it appears that researchers (and, in this case, lawyers) have inadvertently distorted the details of the original study in such a way as to provide a convincing justification for the point they wanted to make.

Moray (1959)

Moray (1959) conducted an early study of selective attention which, as Loftus (1974) has pointed out, has also been subsequently misrepresented in the literature. Moray's experiment consisted of a dichotic listening task where participants were required to repeat out loud a message presented to one ear while other material was being presented in the other ear. Moray discovered that the material presented in the ear that was to be rejected sometimes broke through the attentional barrier. In particular, when the participant's name was presented in the unattended channel and when no prewarnings were given, 33% of the participants reported hearing their name. This experimental finding has become widely known as "the cocktail party phenomenon."

Loftus (1974) cited several authors who incorrectly claimed that Moray found that the participant's name is detected *every* time it is presented in the unattended ear. Note that, in this case, it is an experimental result that is being miscited. The mistake seems to be a typical example of schema-based reconstructive remembering, where the strongest possible result to support the theory (100%) is given in place of the actual result (33%). The error is an important one, however, because these two different results place significantly different constraints on a viable theory of human attention.

Walster (1965)

Walster (1965) conducted an important study of the effects of self-esteem on romantic liking that, according to Rubin (1974), is frequently cited in social psychology textbooks. All of the participants in this experiment were women who were 18 or 19 years old. A short time after the participant arrived for the experiment, a male confederate entered the same room. The confederate struck up a conversation with the participant, and during that conversation, he conveyed personal interest in the participant, and eventually asked her out on a date. Soon after the date was made, the experimenter entered the room, and informed the participant that she would be given a psychological test. The experimenter asked the confederate to take the place of a co-interviewer who supposedly had not yet appeared, and to give the participant the test. Once this test was completed, the experimenter handed the participant either an extremely flattering or an extremely disparaging analysis of a personality test she had taken a few weeks earlier. The experimenter then had the participant fill out a questionnaire asking about her feelings concerning four people, one of whom was the male confederate. The results of the experiment revealed that women whose self-esteem had been temporarily lowered liked the confederate significantly more than did women whose self-esteem had been temporarily raised.

Rubin (1974) has pointed out that several authors citing Walster's (1965) experiment have incorrectly reported the procedure that was adopted. In the published experiment, receipt of approval from the male confederate preceded the self-esteem manipulation but several authors have reversed the order of these two events. Rubin proposes that one of the primary explanations of this error is that the re-ordered sequence seems to provide a better fit to social psychologist's theoretical frameworks than the actual sequence does. So again we find that the details of a well known experiment are reconstructed based on a schema of a canonical experiment supporting a specific theoretical point, thereby leading to predictable distortions. Rubin made a similar observation, comparing the distortions of the Walster experiment to the findings of Allport and Postman (1945).

Schachter (1951)

Schachter (1951) conducted an important investigation of "deviation, rejection, and communication" that, according to Berkowitz (1971), is one of the most widely cited experiments in social psychology. In this study, the participants' degree of attraction to their experimental groups was manipulated by assigning them to clubs whose activities were either more or less preferred. Each club then proceeded to discuss the case history of a juvenile delinquent. The relevance of the discussion topic to each group was also manipulated. In each club, there were three paid participants, one of whom persisted with a deviant opinion throughout the discussion, another who maintained a modal opinion, and a third who began the discussion with a deviate opinion but gradually moved toward the modal view. The primary question of interest was the amount of communication directed towards each of the three paid confederates during the discussion and the participants' attitudes towards these confederates at the end of the session. The results of the experiment are complex and difficult to interpret since the results obtained from various dependent variables were not consistent.

Berkowitz (1971) has pointed out that the results of Schachter's (1951) well-known experiment have been misrepresented by several authors. In his analysis of the citation errors, Berkowitz observed that some authors have misrepresented some of the experimental findings to assimilate them to their own theoretical expectations. Furthermore, he also found that textbook accounts of Schachter's (1951) experiment have omitted several important qualifications, or sometimes have even added details.

The picture that emerges is that authors citing the Schachter (1951) experiment have transformed a complex and somewhat inconsistent set of findings into a clean experimental result, just as one would expect if re-

searchers are reconstructing their recall of the experiment from an idealized experiment schema. Berkowitz himself recognized this connection, citing Bartlett (1932) and Allport and Postman (1945). Once again, the overwhelming tendency to remember a simpler, more consistent depiction of a well-known experiment has led to repeated citation errors.

Discussion

Hull's (1988) theory of science as an evolutionary process has psychological implications that can be tested theoretically and empirically. From a theoretical perspective, we can examine what is known about human remembering to see if it is consistent with the generic properties of an evolutionary process. As discussed earlier, "verb" theories of remembering are consistent with a replication process, analogous to biological evolution. Having established this theoretical correspondence, we can then test for empirical correspondence by examining whether the predictions made from this evolutionary viewpoint are corroborated by the available evidence from the history of science. Six case studies of well-known psychological experiments that had been subject to repeated miscitation errors were collected and reviewed. If science is indeed an evolutionary process with scientists as interactors, scientific beliefs as replicators, and schemata as means for that replication, then we would expect to observe the systematic distortions typical of reconstructive remembering in these six case studies.

Interestingly, the results revealed a consistent pattern of symptoms across the six cases, despite their relatively diverse subject matter. Each of the experiments reviewed is an important and widely cited work. The distortions seem to occur as the findings are disseminated to a larger audience. It is at this point that the process of reconstructive remembering tends to exert its influence. While there is no direct evidence to support the claim that these cases were caused by the properties of human remembering, the pattern of errors that is exhibited across the various cases is uniform and consistent with the "signature" of reconstructive remembering. In all of the cases reviewed, the recall of scientists seems to be driven by an idealized experiment schema that distorts certain details of the actual experiment being reported in predictable ways. Several of the authors who wrote papers to set the historical record straight on these cases of misrepresentation have commented on this fact, pointing out the similarities with the findings of Bartlett (1932) and Allport and Postman (1945).

These systematic distortions typical of reconstructive remembering are indicative of an active, constructive process. They are consistent with the claim that conceptual change in science functions as a selection process with scientists as interactors, scientific beliefs as replicators, and schemata as means for that replication. In short, the descriptive evidence collected

here is consistent with Hull's (1988) account of science as an evolutionary process. This seems to be the first time that psychological evidence of this type has been obtained in favor of Hull's theory.

Despite this contribution, there are several limitations to this work that readily suggest topics for future research. First, it would be useful to assess the generalizability of these results by collecting case studies from disciplines other than psychology. Although there is no reason to believe that these findings are specific to psychology, the conclusions in this article would be bolstered by replications in other areas of science. Second, a more sophisticated test of Hull's (1988) theory would involve an empirical comparison with a competing theory of scientific activity. Ideally, we would identify a phenomenon where the two theories make conflicting predictions in order to pit each theory against the other. Such a comparative experiment would certainly provide stronger evidence than the descriptive findings presented here. Finally, this research has not empirically investigated the conditions under which miscitation errors due to schema-based reconstruction are likely to occur. Known cases of miscitation were deliberately selected, but of course there are many more cases of accurate citation. In addressing this issue, it would be useful to expand the unit of analysis from that of an individual scientist to a group of scientists, thereby providing an opportunity to examine the interaction between the psychological and the social. Of particular interest is the series of findings from group remembering (Hinsz, Tindale, & Vollrath 1997). That research has shown that the recall accuracy of groups is better than that of individuals, one reason being that groups have a better capability for error-detection and error-correction than do individuals. It is likely that this same type of self-correcting process occurs in scientific activity. In fact, Hull's (1988) theory would predict that the competition for the survival of scientific ideas would result in a pressure causing researchers to closely examine the papers of their competitors with the hope of finding flaws, whether they be errors in logic, methodology, empirical correspondence, or attributions. Thus, by investigating the self-correcting properties of remembering in a group of scientists, we have another opportunity to empirically test the theory of science as an evolutionary process.

References

Allport, G. W. & Postman, L. J. 1945. "The basic psychology of rumor." *Transactions of the New York Academy of Sciences* 28: 61–81.

Atkinson, R. C. & Shiffrin, R. M. 1968. "Human memory: A proposed system and its control processes." In K. W. Spence & J. T. Spence Eds.,

The psychology of learning and motivation: Advances in research and theory Vol. 2, pp. 89–195. New York: Academic Press.

Bartlett, F. C. 1932. *Remembering: A study in experimental and social psychology.* Cambridge: Cambridge University Press.

Berkowitz, L. 1971. "Reporting an experiment: A case study in leveling, sharpening, and assimilation." *Journal of Experimental Social Psychology* 7*:* 237–243.

Brewer, W. F. & Nakamura, G. V. 1984. "The nature and functions of schemas." In R. S. Wyer, Jr. and T. K. Srull Eds., *Handbook of social cognition, vol. 1,* pp. 119–160. Hillsdale, NJ: Lawrence Erlbaum.

Caramelli, N. 1987. "The 'schema' concept: Bartlett till now." Paper presented at the Fourth International Conference on Event Perception and Action. Trieste, Italy.

Chase, W. G. & Simon, H. A. 1973*a.* "The mind's eye in chess." In W. G. Chase Ed., *Visual information processing.* pp. 215–281. New York: Academic Press.

Chase, W. G. & Simon, H. A. 1973*b.* "Perception in chess." *Cognitive Psychology* 4: 55–81.

Cornwell, D. & Hobbs, S. 1976. "The strange saga of little Albert." *New Society* 35: 602–604.

de Groot, A. D. 1946. *Het denken van den schaker* [Thought and choice in chess]. Amsterdam: North-Holland.

de Groot, A. D. 1965. *Thought and choice in chess.* The Hague: Mouton.

Harris, B. 1979. "Whatever happened to Little Albert?" *American Psychologist* 34: 151–160.

Hinsz, V. B., Tindale, R. S. & Vollrath, D. A. 1997. "The emerging conceptualization of groups as information processors." *Psychological Bulletin* 121: 43–64.

Hintzman, D. L. 1986. "'Schema abstraction' in a multiple-trace memory model." *Psychological Review* 93: 411–428.

Hull, D. L. 1988. *Science as a process: An evolutionary account of the social and conceptual development of science.* Chicago: University of Chicago Press.

Hull, D. L. 1998. "Studying the study of science scientifically." *Perspectives on Science* 6,3: 209–231.

Jongman, R. W. 1968. *Het oog van de meester* [The eye of the master]. Amsterdam: van Gorcum.

Kelso, J. A. S. 1995. *Dynamic Patterns: The self-organization of brain and behavior.* Cambridge, MA: MIT Press.

Kolers, P. A., & Roediger, H. L., III 1984. "Procedures of mind." *Journal of Verbal Learning and Verbal Behavior* 23: 425–449.

Loftus, E. F. 1974. "On reading the fine print." *Quarterly Journal of Experimental Psychology* 27: 324.

Logan, G. D. 1988. "Toward an instance theory of automatization." *Psychological Review* 95: 492–527.

Moray, N. 1959. "Attention in dichotic listening: Affective cues and the influence of instructions." *Quarterly Journal of Experimental Psychology* 11: 56–60.

Morris, C. D., Bransford, J. D., & Franks, J. J. 1977. "Levels of processing versus transfer appropriate processing." *Journal of Verbal Learning and Verbal Behavior* 16: 519–533.

Neisser, U. 1982. *Memory Observed: Remembering in natural contexts.* San Francisco: Freeman.

Neisser, U. & Winograd, E. 1988. *Remembering Reconsidered: Ecological and traditional approaches to the study of memory.* Cambridge: Cambridge University Press.

Plotkin, H. 1994. *Darwin Machines and the Nature of Knowledge: Concerning adaptations, instinct and the evolution of intelligence.* London: Penguin.

Prytula, R. E., Oster, G. D., & Davis, S. F. 1976. "The 'rat rabbit' problem: What did John B. Watson really do?" *Teaching of Psychology* 4: 44–46.

Rubin, Z. 1974. "A new order of things: A case study in the reconstruction of an experimental procedure." *Representative Research in Social Psychology* 5: 81–87.

Schachter, S. 1951. "Deviation, rejection, and communication." *Journal of Abnormal and Social Psychology* 46: 190–207.

Schacter, D. L., Harbluk, J. L. & McLachlan, D. R. 1984. "Retrieval without recollection: An experimental analysis of source amnesia." *Journal of Verbal Learning and Verbal Behavior* 23: 593–611.

Simonton, D. K. 1997. "Creative productivity: A predictive and explanatory model of career trajectories and landmarks." *Psychological Review* 104: 66–89.

Thelen, E. & Smith, L. B. 1994. *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

Treadway, M. & McCloskey, M. 1987. "Cite unseen: Distortions of the Allport and Postman rumor study in the eyewitness testimony literature." *Law and Human Behavior* 11: 19–25.

Vicente, K. J. & de Groot, A. D. 1990. "The memory recall paradigm: Setting the historical record straight." *American Psychologist* 45: 285–287.

Walster, E. 1965. "The effect of self-esteem on romantic liking." *Journal of Experimental Social Psychology* 1: 184–197.

Watson, J. B. & Rayner, R. 1920. "Conditioned emotional reactions." *Journal of Experimental Psychology* 3: 1–15.

Waugh, N. C. & Norman, D. A. 1965. "Primary memory." *Psychological Review* 80: 352–373.