

Available online at www.sciencedirect.com



Journal of Memory and Language 55 (2006) 441-446

Journal of Memory and Language

www.elsevier.com/locate/jml

Some targets for memory models *

Stephan Lewandowsky a,*, Evan Heit b

a School of Psychology, University of Western Australia, 35 Stirling Highway, Mail Bag M304, Crawley, W.A. 6009, Australia
 b University of California, Merced, USA

Received 9 August 2006

Abstract

This introductory article to the *Journal of Memory and Language* special issue on memory models discusses the progress made in the field of memory modeling during the last few decades in terms of a number of previously suggested criteria, using the articles in this issue as examples. There has been considerable progress, both at a technical level (e.g., concerning model comparison and model analysis techniques) and at a psychological level (as evidenced by the increasingly tight interplay between theory and data on human memory). The article concludes by proposing a few generic targets for future modeling work.

© 2006 Elsevier Inc. All rights reserved.

Perhaps more so than in other branches of experimental psychology, the progress made in the mathematical modeling of memory during the last few decades has been phenomenal, at least if measured by the sheer quantity of output and by the increasing sophistication of computational techniques. The articles that appear in this special issue of the *Journal of Memory and Language* represent the leading edge of current progress in this burgeoning field.

At a time when there are important concerns about how best to analyze data from psychology experiments (e.g., Loftus, 2001), it is particularly appropriate to highlight the role of mathematical and computational mod-

eling. Our intention in this introductory article is twofold: First, we reflect on the extent of scientific progress achieved through memory modeling during the last few decades and examine how the papers in this issue relate to that progress. Second, we briefly offer our own thoughts on what targets remain to be accomplished in future research.

Progress in computational modeling of memory

Turning first to an examination of the somewhat vexing issue of what exactly constitutes "progress," one cannot help but note that previous discussions of the issue were sometimes characterized by angst about whether the field as a whole has made sufficient cumulative advances (e.g., Krantz, Atkinson, Luce, & Suppes, 1974, cited in Estes, 1975). We take as our lead a discussion of the issue of progress and its proper metric in an influential article by Estes (1975), titled "Some targets for mathematical psychology." Estes touched on a large number of issues relating to the fate of mathematical

[†] The authors are grateful to Kay Bock, Bob Greene, and Dave Plaut for comments on this article. Preparation of this article was facilitated by a Large Grant and by a Discovery Grant from the Australian Research Council to the first author.

^{*} Corresponding author. Fax: +61 8 6488 1006.

E-mail address: lewan@psy.uwa.edu.au (S. Lewandowsky). *URL:* http://www.psy.uwa.edu.au/user/lewan/ (S. Lewandowsky).

and computational modeling, among them the disappointing inability of mathematical psychology to contribute to the solution of social problems. Although this disappointment seems to have abated, possibly reflecting a gradual resolution of those problems or perhaps declining interest in them, three other issues and suggestions raised by Estes remain pressing to this date.

(1) A core enterprise of computational modeling involves making a choice between different competing models, in order to select the one that best explains the data. Almost inevitably, one central aspect of this process—although by no means the only one—consists of the evaluation of the relative fit of the various models to data. The details of this comparative evaluation are, however, far from trivial and Estes (1975) bemoaned the "extreme meagerness" (p. 270) of existing techniques to permit proper model comparison.

Fortunately, the intervening three decades have seen tremendous progress in model comparison techniques, and it is now commonplace to see various different models compared by formal statistical tests based on maximum likelihood estimation (see Myung, 2003; for a tutorial introduction), techniques that in the 1970s were not widely available. Indeed, the sophistication of model comparison has increased to the point where a number of fit statistics exist that penalize models for parametric complexity, thus permitting comparison between the proverbial apples and oranges (e.g., Myung, 2000). Accordingly, several of the papers in this issue (e.g., by Heathcote, Raymond, & Dunn and Oberauer & Kliegl, Rotello & MacMillan) have used model comparison techniques that seek to correct for differences in complexity among the various different models.

Readers interested in more details on model selection may wish to consult a recent special issue of the *Journal of Mathematical Psychology* (Volume 50, Issue 2, April 2006) that was dedicated to the topic and that reveals the stunning technical progress made during the last few decades.

Notwithstanding the obvious requirement that models have to fit the data they seek to explain, Estes (1975) additionally argued that "...intensive efforts should be directed toward finding additional bases for evaluating quantitative theories" (p. 268). Again, as in the case of model selection, there has been much progress towards addressing this concern. For example, Li, Lewandowsky, and DeBrunner (1996) proposed that the utility of a model can be examined by analyzing the behavior of its parameters, and that inferences about the scope and hence falsifiability-of a model can be made on the basis of that analysis, even before further data become available to challenge the model. Another proposal involves examining how well one model fits data generated by an alternative model B (and vice versa), a technique known as "landscaping" (Navarro, Pitt, & Myung, 2004), which has been shown to hold considerable promise. Inspection of the articles in this special issue reveals that those techniques have yet to be widely adopted by memory researchers, and their increased use forms a desirable target for future modeling work in human memory to which we return below.

Overall, there is no doubt that our tools for the selection and evaluation of computational models have become far better than was perhaps even imaginable in the 1970s. The question, then, is whether these technical advances have also advanced our psychological understanding of memory.

(2) Estes (1975) proposed that one indisputed utility of models was to permit the classification of data into those results that are fit by a certain model versus those that are not. Far from presenting a problem (because the model then necessarily mispredicts at least some of the data), Estes suggests that such classification, akin to how physicists view light as a wave to explain some results and as a particle to explain others, might be preferable to attempts to build more encompassing theories which, despite handling more data, represent an unwelcome increase in complexity.

In contrast to the clear technical advances that the field has delivered, there is less evidence that Estes's admonition to categorization of data has been heeded. On the contrary, the trend has been towards integrative theories, but within circumscribed domains of research. For example, we now have theories of recognition memory (McClelland & Chappell, 1998; Shiffrin & Steyvers, 1997) or of short-term memory (e.g., Brown, Neath, & Chater, 2002; Burgess & Hitch, 1992, 1999; Farrell & Lewandowsky, 2002) that account for much of the available data and that have stimulated much additional experimentation. Accordingly, predictions of the preceding models were examined in the papers by Criss (this issue), and by Surprenant, Neath, and Brown, by Burgess and Hitch, and by Farrell (all in this issue), respectively. However, notwithstanding the repeated tests of those models, within their particular domain, there appears to be little systematic categorization of phenomena into those that one theory can handle versus those that are best accommodated by another.

By contrast, between domains, there appears to be a strict and largely unintentional categorization that is quite unlike that intended by Estes (1975). For example, theories of short-term memory have, until recently, existed in a virtual bubble, reminiscent of Rowan Atkinson's "Mr. Bean," with little or no connection to closely related areas of research such as working memory or long-term memory. Indeed, the nearly complete separation between two entire bodies of research and theorizing according to whether people remember a list of words (i.e., a "short-term" memory task) or a list consisting of the final words of sentences (i.e., a "working" memory task) surely cannot be taken to be a sign of healthy classification. Likewise, the nearly exhaustive

consideration of short-term memory tasks by many theories, without any acknowledgement of the role of longterm learning and memory, is unlikely to constitute the classification of phenomena intended by Estes (1975).

We would add that this unfortunate compartmentalization not only separates one area of memory research from another, but also separates memory research from other areas of experimental psychology. Exceptions to this compartmentalization involve Estes's own later work (1994), which used the same modeling framework to address both recognition memory and categorization, while also dealing with recall, short-term memory, and simple reasoning tasks. Other examples of modeling intended to make connections between different areas of experimental psychology include, for memory and categorization, Nosofsky (1988) and Nosofsky and Zaki (1998), and for memory and reasoning, Heit and Hayes (2005) and Heit and Rotello (2005). Furthermore, the work by John R. Anderson (e.g., Anderson & Lebiere, 1998) has culminated in a comprehensive modeling approach that bridges virtually all areas of cognition.

Likewise, some of the articles in this special issue are intended to avoid compartmentalization and make connections between areas that have been usually been studied in isolation: The paper by Burgess and Hitch (this issue) models the intriguing interactions between longterm memory and short-term retention; and the paper by Oberauer and Kliegl (this issue) hints at the connection between short-term and working-memory tasks. Similarly, the paper by Kent and Lamberts (this issue) draws welcome connections between perception and recognition, and Berry, Henson, and Shanks (this issue) unify priming and recognition under a common theoretical umbrella. Finally, Jones, Kintsch, and Mewhort (this issue) relate priming to semantic and episodic memory. We return to the issue of bridging research areas below, when we propose our own targets for future computational modeling work in human memory.

(3) To illustrate some of the preceding points, Estes (1975) presented a case study of theory development in paired-associate learning. The details of this case are of little contemporary interest, but it illustrates the clear historical shift from single diagnostic experiments that mandate identifiable and tractable modification of a model (as reviewed by Estes, 1975) to a vastly greater enterprise in which theories now have to handle a very large set of extant data to be taken seriously (e.g., any plausible model of serial recall must, at the very least, handle the extensive primacy and small recency of the serial position curve; the symmetry of transposition error gradients; the increase of omission and intrusion errors across serial positions; and a sometimes frustratingly large number of additional effects). Perhaps as a consequence of this requirement to fit a larger data base, current models often come in different variants (e.g., the model by Burgess & Hitch has been presented in three different variants; 1992, 1999, and in this issue) or can change their theoretical nature dramatically as a function of parameter settings (e.g., a time-based theory can turn into an event-based theory with the change of a single parameter value; Lewandowsky, Duncan, & Brown, 2004).

The problem of theory flexibility and model mimicry (i.e., that one model can behave like another one, hence preventing discrimination) has been discussed quite extensively in the model selection literature (e.g., Myung, 2000; Pitt, Kim, & Myung, 2003), and gets some attention in the current issue from Criss and McClelland. We return to the problem of flexibility below, when we postulate our own targets for model development.

Although Estes's (1975) paper has been deservedly influential, his criteria for progress are not all-encompassing. An alternative approach to the question of progress was taken by Murdock (1974). Virtually contemporaneously with Estes (1975), Murdock (1974) cited the interrelationship between theory and data as a critical element of progress: "To segregate theory and data is not a very satisfactory procedure. They should be thoroughly and carefully interrelated. However we are not quite ready for that. On the one hand, none of the theories we now have explain in any depth and with great precision even an appreciable fraction of the relevant data. On the other hand, to consider the data alone would be barren and uninteresting" (Murdock, 1974, p. xi).

Using the interrelationship between data and theory as a criterion, Lewandowsky and Hockley (1991) examined progress during the 1970s and 1980s in three specific arenas of memory research; item recognition, serialorder memory, and memory for associative information. They concluded that considerable progress had indeed been made in all arenas, a claim best illustrated here with respect to item recognition. On the theoretical front, the period from the mid-1970s onward has been characterized by the abandonment of a simple scanning notion and the migration towards more complex models that are best understood by considering the papers by Criss, Rotello and Macmillan; and Heathcote, Raymond, and Dunn in this issue. The impetus for this theoretical development was provided by the everincreasing empirical precision and the development of incisive new experimental measures during the same time period. For example, the shift in empirical focus from mean latencies to analysis of the distribution of latencies (Ratcliff & Murdock, 1976) and the advent of the response-deadline methodology (e.g., Gronlund & Ratcliff, 1989; see Kent & Lamberts and Oberauer & Kliegl, this issue, for recent applications) strongly constrained theory development and contributed to the demise of serial scanning models (e.g., Hockley, 1984). In consequence, the field now enjoys a complex and

mutually constraining inter-twining of theory and data that represents a radical and qualitative advance over the situation 30 years ago, when the apparent linearity of set size functions gave rise to the notion of serial scanning.

However, notwithstanding this generally rosy picture, we are concerned that the field may regress towards a less sophisticated empirical pose. We take up this issue below, when postulating our targets for memory models.

Finally, no discussion of progress can be complete without considering what is perhaps the most exciting current development in the field that, to our knowledge, no previous discussion of progress in the 1970s or 1990s had anticipated. The rising tide of Bayesian approaches to modelling of memory and cognition is nothing short of revolutionary, as a glimpse at the table of contents of the most recent *Proceedings* of the meetings of the Cognitive Science Society will confirm. Readers interested in the Bayesian approach, which links human behavior to normative statistical expectations, may also wish to consult a recent special issue of Trends in Cognitive Sciences (Volume 10, Issue 7, July 2006). The papers by Criss in this issue are at least partially inspired by the Bayesian approach, thus providing a close-up glimpse at the Bayesian "revolution."

The excitement surrounding the Bayesian approach mirrors the excitement and enthusiasm that, some two decades ago, met the connectionist "revolution" (Rumelhart et al., 1986). More distantly, the import of signal detection theory into memory research may have caused a similar stir another decade or so earlier (cf. Estes, 1975). It is informative to consider what has happened to connectionism (and signal detection theory) in the meantime. In both cases, the approach has been transformed from an awe-inspiring panacea to a wellfounded and well-understood tool that continues to contribute to our understanding of memory performance. For example, in the current issue, the papers by Rotello and Macmillan; Berry, Henson, and Shanks; and Heathcote, Raymond, and Dunn constitute leading-edge applications of signal-detection theory, whereas the papers by Farrell and by Burgess and Hitch advance a connectionist metaphor for the processes they are examining. We suggest that Bayesian approaches to cognition will similarly turn into additional powerful arrows in the psychologists' quiver that will facilitate our future understanding of human memory.

Some targets for memory modeling

Our analysis leads us to postulate four targets for future endeavors in the modeling of human memory. Although those targets are necessarily subjective, we keep them as generic as possible by avoiding focus on particular domains of enquiry.

- (1) Overcoming "isolationism." We have already pointed out how some of the papers in this issue are beginning to build bridges between areas of enquiry. This trend is welcome and must continue if we wish to overcome some of the rather arbitrary divisions discussed earlier. Our rank ordering of the urgency with which these divisions should be addressed places the division between research on short-term memory and working memory at the top of the list. None of the short-term memory models considered in this issue (i.e., in the papers by Burgess & Hitch and by Farrell) can be applied to working-memory tasks, and conversely, the modeling of working memory by Oberauer and Kliegl (this issue) has little to say about the modal results in short-term memory. Given that both areas of enquiry share crucial methodological and conceptual features, this separation—and the need to overcome it—is striking. Other divisions, however, are more apparent than real and their linkage may safely be postponed. For example, much of the current modeling of recognition memory is performed without apparent connection to the parallel literature on recall or short-term memory. However, at least in the case of the Bayesianinspired REM framework (e.g., see the papers by Criss in this issue), this represents a strategic choice because the underlying theoretical framework embraces most facets of memory (Shiffrin, 2003), thus ensuring longterm convergence of modeling efforts that are currently conducted in parallel.
- (2) Resisting "slippage." Much of the progress in recent years has been due to the development of sophisticated measures and techniques, such as the use of response-deadline methodologies or the analysis of response-time distributions. Concerning the latter, we find it noteworthy that after the use of reaction-time distributions had been pioneered in research on recognition memory, current work again appears to rely more on simple latency and accuracy measures. None of the papers in this issue consider response—time distributions, and it would be unfortunate indeed if this were indicative of a trend towards diminishing empirical constraints on modeling.
- (3) Importing better model comparison techniques. The literature on model selection contains much new information about techniques for model comparison and evaluation, ranging from maximum-likelihood methods of model comparison to "landscaping," parameter sensitivity analysis, and many more that we did not have room to discuss. Future modeling work would benefit from use of those additional techniques, in particular in the context of examining the flexibility of current models.
- (4) Accounting for individual differences. This final target has not been anticipated by previous analyses of progress in memory modeling; it also has little prominence in the model selection literature. Nonetheless, we

believe that modeling of individual differences in memory must be a target for future research.

One reason is simply practical: Averaging data across participants necessarily leads to a loss of data and sometimes even obscures the patterns shown by individuals (e.g., Estes & Maddox, 2005). Second, the pattern of individual differences may be as revealing as the behavior of the data in the aggregate. A good precedent for this claim can be found in categorization research, where individuals' generalization profiles (e.g., Nosofsky, Palmeri, & McKinley, 1994) or at least the analysis and modeling of subgroups of participants (e.g., Yang & Lewandowsky, 2004) are now fairly routine and have arguably contributed to our advanced understanding of categorization.

Examples in this special issue of fitting individual participants' data include the articles by Kent and Lamberts and Heathcote, Raymond, and Dunn. The article by Oberauer and Kliegl showcases a particularly elegant way of analyzing both group and individual data, in terms of non-linear mixed effects models. In addition, Surprenant, Neath, and Brown (this issue) compare short-term memory in older and younger adults. These articles are—one hopes—harbingers of future developments in this area.

On this note, what would be most valuable is not only modeling of individuals but also having a coherent account of how people differ. For example, one issue in working memory research has been explaining correlations between IQ and various facets of (working) memory span (e.g., Conway, Kane, & Engle, 2003). As in other areas of psychological research (Stanovich, 1999), theoretical development in memory research could be spurred by broader and more systematic study of which memory components (beyond working memory) are or are not correlated with IQ. Most important in the present context, the study of those correlations should be accompanied by computational modeling of the underlying cognitive processes, something that has been virtually completely absent to date. Of course, this call is tantamount to a call for the development of a complete computational process model of not just memory but also intelligence itself—clearly a big task whose completion time will be measured in decades, not years. Lest one think that this goal is too ambitious to be even considered at this moment in scientific history, the reader may wish to recall that maximum likelihood techniques appeared to be a distant dream only three decades ago (Estes, 1975).

Finally, although we are hesitant to nominate the resolution of social problems as another target for memory modeling, we find it self-evident that memory theorists are not operating in a societal vacuum and that connections between abstract theory and big societal issues can—and indeed should—be drawn. As a case in point, consider the memory updating paradigm investigated by

Oberauer and Kliegl (this issue). The ability to update working memory, and to discard information that is no longer needed, arguably forms a crucial aspect of text comprehension, in particular when people must process corrections of earlier information (e.g., when a fire that is initially blamed on an arsonist turns out to be due to an electrical fault: cf. Johnson & Seifert, 1994). The way in which people process corrections in the laboratory, in turn, has been identified also to govern people's persistent misperceptions of "real" news stories, such as the continued widespread belief that Weapons of Mass Destruction were found in Iraq after the 2003 invasions (Lewandowsky, Stritzke, Oberauer, & Morales, 2005). Computational models of memory updating may therefore eventually make an important contribution to some very big issues indeed.

Conclusions

In our view, the articles compiled in this special issue indisputably reveal that modeling of human memory has made tremendous progress during the last 30 years. Although much remains to be done—and we have postulated some tentative targets for future endeavors—the abundance of novel technical tools and the accompanying increased precision of our theories and sophistication of the data base undeniably represent real progress. The self-criticism and skepticism that characterized earlier stock-takings now appears misplaced: The sum total of our knowledge about human memory differs qualitatively from anything that was available at the time of Estes (1975) or even Lewandowsky and Hockley (1991).

References

Anderson, J. R., & Lebiere, C. (1998). The atomic components of thought. Mahwah, NJ: Erlbaum.

Brown, G.D.A., Neath, I., Chater, N. (2002). SIMPLE: A local distinctiveness model of scale invariant memory and perceptual identification. Unpublished manuscript.

Burgess, N., & Hitch, G. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language*, 31, 429–460.

Burgess, N., & Hitch, G. (1999). Memory for serial order: a network model of the phonological loop and its timing. *Psychological Review*, 106, 551–581.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Science*, 7, 547–552.

Estes, W. K. (1975). Some targets for mathematical psychology. *Journal of Mathematical Psychology*, 12, 263–282.

Estes, W. K. (1994). Classification and cognition. New York/ Oxford: Oxford University Press.

Estes, W. K., & Maddox, W. T. (2005). Risks of drawing inferences about cognitive processes from model fits to

- individual versus average performance. Psychonomic Bulletin & Review, 12, 403-408.
- Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, 9, 59–79.
- Gronlund, S. D., & Ratcliff, R. (1989). Time course of item and associative information: implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 846–858.
- Heit, E., & Hayes, B. K. (2005). Relations among categorization, induction, recognition, and similarity. *Journal of Experimental Psychology: General*, 134, 596–605.
- Heit, E., Rotello, C.M. (2005). Are there two kinds of reasoning? Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society, 923–928. Mahwah, NJ: Erlbaum.
- Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 598–615.
- Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: when misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20, 1420–1436.
- Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004).
 Time does not cause forgetting in short-term serial recall.
 Psychonomic Bulletin & Review, 11, 771–790.
- Lewandowsky, S., & Hockley, W. E. (1991). Relating theory and data: Towards an integration. In W. E. Hockley & S. Lewandowsky (Eds.), Relating theory and data: Essays on human memory in honor of Bennet B. Murdock (pp. 3–20). Hillsdale, NJ: Lawrence Erlbaum.
- Lewandowsky, S., Stritzke, W. G. K., Oberauer, K., & Morales, M. (2005). Memory for fact, fiction, and misinformation: the Iraq War 2003. *Psychological Science*, 16, 190–195.
- Li, S.-C., Lewandowsky, S., & DeBrunner, V. E. (1996). Using parameter sensitivity and interdependence to predict model scope and falsifiability. *Journal of Experimental Psychology: General*, 125, 360–369.
- Loftus, G. R. (2001). Analysis, interpretation, and visual presentation of data. Stevens' handbook of experimental psychology (3rd ed.). New York: John Wiley.

- McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: a subjective likelihood approach to the effects of experience in recognition memory. *Psychological Review*, 105, 734–760.
- Murdock, B. B. (1974). *Human memory: Theory and data*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Myung, I. J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology*, 44, 190–204.
- Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, 47, 90–100.
- Navarro, D. J., Pitt, M. A., & Myung, I. J. (2004). Assessing the distinguishability of models and the informativeness of data. *Cognitive Psychology*, 49, 47–84.
- Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 700–708.
- Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994).Rule-plus-exception model of classification learning. *Psychological Review*, 101, 53–79.
- Nosofsky, R. M., & Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: an exemplar-based interpretation. *Psychological Science*, 9, 247–255.
- Pitt, M. A., Kim, W., & Myung, I. J. (2003). Flexibility vs. generalizability in model selection. *Psychonomic Bulletin & Review*. 10, 29–44.
- Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review*, 83, 190–214.
- Rumelhart, D.E., McClelland, J.L., and the PDP Research Group. (1986). Parallel distributed processing: Explorations in the microstructure of cognition, vols. 1 and 2. Cambridge, MA: MIT Press.
- Shiffrin, R. M. (2003). Modeling memory and perception. *Cognitive Science*, 27, 341–378.
- Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—Retrieving effectively from memory. Psychonomic Bulletin & Review, 4, 145–166.
- Stanovich, K. E. (1999). Who is rational? Studies of individual differences in reasoning. Mahwah, NJ: Erlbaum.
- Yang, L.-X., & Lewandowsky, S. (2004). Knowledge partitioning in categorization: constraints on exemplar models. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, 1045–1064.