

Original Article

The Difficulty with Experience: Does Practice Increase Susceptibility to Premature Closure?

Kevin W. Eva, PhD, and John P. W. Cunningham, MD

Abstract

Introduction: A recent review of the physician performance literature concluded that the risk of prematurely closing one's diagnostic search increases with years of experience. To minimize confounding variables and gain insight into cognitive issues relevant to continuing education, the current study was performed to test this conclusion.

Methods: Physician participants were shown a series of case histories and asked to judge the probability of a pair of diagnoses. The order in which features were presented was manipulated across participants and the probabilities compared to determine the impact of information order. Two groups of participants were recruited, 1 older than and 1 younger than 60 years.

Results: The probability assigned to a diagnosis tended to be greater when features consistent with that diagnosis preceded those consistent with an alternative than when the same features followed those consistent with the alternative. Older participants revealed a greater primacy effect than less experienced participants across 4 experimental conditions.

Discussion: Physicians with greater experience appear to weigh their first impressions more heavily than those with less experience. Educators should design instructional activities that account for experience-specific cognitive tendencies.

Key Words: Clinical reasoning, aging, clinical experience, nonanalytic processes, information order, continuing education, medical

In North America and worldwide, we are experiencing demographic drift toward an older population. Whereas a great deal of effort has been expended attempting to predict the impact this drift will have on the health care requirements of society, attention has only recently turned to the

effects aging will have on the performance of the health care providers themselves. Already, however, a fourth to a third of physicians are older than 55 years,¹ making it important for the successful design of continuing education efforts to develop an understanding of the effects of age-related cognitive tendencies on the maintenance of expertise.

In a systematic review published in 2005, Choudhry and colleagues found that 45 of 62 studies (73%) evaluating the relation between years of experience and quality of health care showed some degree of decreasing performance with increasing experience whereas 13 (21%) reported no association.² Their analysis makes a compelling argument for the tendency of physicians to lose track of the state of the art because many of the outcomes focused on formal knowledge, but there is an aspect of expertise that

Dr. Eva: Department of Clinical Epidemiology and Biostatistics, Program for Educational Research and Development, McMaster University;

Dr. Cunningham: Department of Medicine, McMaster University, and Physician Review and Enhancement Program, CPSO, Hamilton, Ontario.

Partial financial support for this project was provided by the Society for Academic Continuing Medical Education.

Correspondence: Kevin W. Eva, PhD, Department of Clinical Epidemiology and Biostatistics, Program for Educational Research and Development, MDCL 3522, McMaster University, 1200 Main St. W., Hamilton, Ontario, Canada L8N 3Z5; e-mail: evakw@mcmaster.ca.

cannot be accounted for by many of the measures utilized in their review: experience.³ That is, experience provides something that cannot be gained from being able to quote the most recent practice guidelines. For example, Hobus et al.⁴ reported a strong positive correlation between competence and years of experience when the clinical information provided to physicians was limited to the contextual information that would be obtained early in a patient encounter. This, combined with a large literature on the benefits of reasoning through pattern recognition, suggests an intriguing paradox. Physicians who have been in practice longer appear less likely to keep up-to-date, yet they appear to have more efficient diagnostic skill.

In reviewing the psychology literature, we have noted that many laboratory studies of cognitive tendencies suggest a mechanism whereby such a paradox might arise.^{5,6} Common in the literature on human memory is the notion that both analytic (ie, controlled) and nonanalytic (automatic) mental processes contribute to performance. One can accurately conclude that a stimulus has been seen previously either by true and explicit recall of the past experience or through more nonanalytic processes including the use of familiarity. Jacoby and colleagues^{7,8} have shown over a series of experiments that the memory failures associated with presenile aging tend to be failures of analytic control whereas automatic memory processes tend to remain intact across the life span.

The distinction between analytic and non-analytic processes has formed a similarly important divide between models of diagnostic reasoning.⁹ Stated plainly, diagnostic decision making involves attempting to categorize case histories as indicative of specific diagnoses. One can complete such a task analytically by gathering information (including base rates and the presence or absence of signs and symptoms) so that the relative probabilities of multiple diagnoses can be carefully considered, adjusted, and weighed against one another. In

recent decades, however, the medical and psychological communities have come to recognize that nonanalytic processes also play a substantial role in the diagnosis of clinical cases.^{10,11} Accurate diagnoses can be made based on the recognition (often unconscious) that the current case is similar to one that has been seen in the past. In fact, despite the admonitions of medical educators to “be objective and systematic,” when instructional manipulations have been used to compare the strength of analytic and nonanalytic approaches to diagnosis, attempting to adopt a purely analytic strategy has been shown to impede performance relative to when nonanalytic strategies are advocated as useful supplements to the diagnostic process.^{12,13}

To date, however, most models of diagnostic expertise in medicine have not incorporated dynamic contributions of both types of processing.^{9,14} Nor have they suggested that the relative contributions of both types of processing might change during the generation and maintenance of expertise. In evaluating the literature on physician performance, we have noted that the most parsimonious explanation for the existing pattern of data is one analogous to that of Jacoby’s work on human memory: an age-related decline in the analytic contribution to diagnostic reasoning accompanied by stability in the nonanalytic contribution.⁵ If the initial hypotheses raised during a categorization task such as medical diagnosis tend to arise from automatic, nonanalytic processes, then a greater amount of experience should improve physicians’ ability to generate plausible diagnoses during the early stages of a patient encounter. However, if confirmation is more controlled and analytic and the controlled aspects of memory tend to decrease with aging, it is plausible that more experienced physicians might become less likely to retain conflicting details of case histories long enough to allow them to overrule their initial conception of clinical cases. Indeed, careful examination of data from physician review programs suggests that

troubled older physicians tend to have difficulty as a result of prematurely closing their diagnostic search; the most prevalent errors include non-comprehensive history taking, incomplete data gathering, and failure to consider important management strategies.¹⁵

Purpose

Although both laboratory studies of memory and real-world data from physicians are supportive of this model, an experimental study that used clinical material was deemed necessary to (1) rule out the role of potentially confounding variables (eg, differences in practice patterns) that affect the somewhat naturalistic data that focus specifically on physician performance, and (2) to lessen concerns about ecological validity that inevitably arise when generalizing from studies of word-list recall such as those found in the psychology literature referred to above. This was the goal of the current research project.

Method

To begin studying this issue, we adapted an “information order” protocol, asking physicians to diagnose a case history in which the order of feature presentation was manipulated. Cunningham et al.¹⁶ presented diagnosticians with case histories in which diagnostic features consistent with pneumonia, for example, were presented either before or after diagnostic features consistent with pulmonary embolus. A primacy effect was observed in that pneumonia was considered to be more likely when the features consistent with pneumonia were presented early in the case history than when the same features were presented late. The age or years of experience of the physicians was, unfortunately, not recorded in that study. Given the proposed model of expert performance and the literature that suggests that nonanalytic processes act rapidly, often without conscious awareness, we anticipated a larger primacy effect (ie, an increased reliance on first

impressions) among older physicians than among younger physicians.

Ethics approval was granted by the Hamilton Health Sciences Research Ethics Board. Fifteen family physicians, 7 younger than 60 and 8 60 or older, participated. Participants were asked to read 8 clinical cases with the goal of determining the correct diagnosis. Each case was created with 2 competing diagnoses in mind. Across the 8 cases, the diagnoses were

1. Pneumonia versus pulmonary embolus
2. Lung cancer versus bronchitis
3. Migraine versus subarachnoid hemorrhage
4. Myocardial infarction versus gastritis
5. Infectious disorder versus inflammatory bowel disease
6. Peptic ulcer versus gallstones
7. Gout versus infectious arthritis
8. Appendicitis versus pelvic inflammatory disorder

Across participants, each case was presented in 1 of 2 orders: The features consistent with diagnosis A were presented before or after the features consistent with diagnosis B. For example, case 2 was presented to 8 participants with the features consistent with lung cancer preceding those consistent with bronchitis, and the reverse was true for the remaining physicians. Participants were asked after each case presentation to assign probability judgments (using a scale of 0 to 100) indicating the likelihood that the patient suffered from each of 2 diagnostic possibilities. Comparison of the probability ratings assigned across the 2 orders determined the magnitude of any primacy effects for each case. Four experimental conditions were used for peripheral reasons that cannot be fully explained because of space constraints but will be described briefly in the discussion. Each participant viewed 2 cases within each condition, the order in which conditions were presented being counterbalanced across participant:

- *Generated:* In this condition, participants were presented with cases that were nearly equally indicative of 2 diagnoses. Participants were asked to generate diagnostic possibilities and then assign probability ratings.

- *Provided:* The same as the generated condition except diagnoses were provided. Participants merely had to assign probability ratings.

- *Privileged:* The same as the provided condition except an additional piece of very important diagnostic information for 1 of the 2 diagnoses (eg, a past history of cancer being indicative of lung cancer in case 2) was manipulated (ie, presented either early or late in the case presentation).

- *Extreme:* The same as the privileged condition except that features were added and deleted from the cases such that the diagnosis counterindicated by the privileged piece of information (eg, bronchitis in case 2) would be viewed as substantially more probable than the diagnosis indicated by the privileged piece of information. In other words, bronchitis and lung cancer were no longer roughly equiprobable when case 2 was presented in the extreme condition: the features were manipulated such that bronchitis became substantially more likely.

Performance time was not measured, but explanation of the procedure, data collection, and debriefing required approximately 50 minutes per participant. Each case (and the accompanying experimental instructions) was presented by laptop computer, typically in the participants' place of employment. The cases were presented sequentially, 1 sentence per slide, and responses were collected using pen and paper.

Results

The mean age of the younger and older cohorts of physicians was 54.1 years (SD 4.12) and 71.6 years (SD 1.42), respectively. As could be anticipated, the number of years practicing as a family physician correlated highly with age ($r = 0.95$),

the means for the 2 groups being 22.3 years (SD 6.48) and 42.7 years (SD 2.74), respectively.

A mixed design analysis of variance was performed on the probability judgments, with order (defined as the diagnosis for which features were presented early versus the diagnosis for which features were presented late) being treated as a repeated measure. Age group (younger than 60 v 60 or older) and condition (generated, provided, privileged, and extreme) were treated as between-subject factors. In no analyses did "case" reveal a significant main effect or a significant interaction with other variables. Therefore, it will be ignored in the reporting of these results. Overall, a primacy effect was observed, diagnoses receiving higher probability ratings when their indicative features were presented in the case earlier (mean, 52.1; 95% confidence interval [CI], 47.0–57.3) than when the same features were presented later (mean, 37.8; 95% CI, 33.3–42.3; $p < .01$).

The magnitude of this effect interacted with age group, younger physicians revealing a smaller primacy effect (mean = 47.3 when features were presented early relative to 43.5 when features were presented late, primacy effect = 3.8) than older physicians (means, 57.0 and 32.1; primacy effect, 24.9; $p < .03$).

In addition, the effect of information order was variable across conditions as indicated by a statistically significant order times condition interaction ($p < .05$). These differences are illustrated in Table 1. No other main effects or interactions were statistically significant ($p > .50$ in all cases).

Discussion

This study aimed to reconcile the discrepancy that has been observed between comprehensive assessments of physician performance (that tend to reveal a negative relationship with age) and diagnostic tasks in which only contextual information is provided (and tend to reveal a positive relationship with age). More specifically, we

Table 1 Mean Probability Rating (and Difference Scores) as a Function of Condition and Cohort

Condition	Order	Less Than 60 Years		60 Years and Older	
		Probability Assigned	Difference	Probability Assigned	Difference
Generated	First	57	25	68	46
	Second	32		22	
Provided	First	48	10	57	27
	Second	38		30	
Privileged	First	48	4	58	27
	Second	44		33	
Extreme	First	35	-23	43	2
	Second	58		41	

have attempted to advance scientific understanding of the relationship between age, expertise, and reasoning processes by testing the sufficiency of a model of expert categorization that we believe to be capable of resolving this paradox.^{5,9} The model consists of 2 interactive processes, 1 being primarily nonanalytic and 1 being primarily analytic. Our position is that performance is maximized during categorization tasks by forming a nonanalytic case representation, based on similarity to past cases, for example, *and* engaging in more analytic confirmation to ensure that all plausible diagnoses are given the appropriate amount of consideration. Older physicians, it was hypothesized, would tend to show a decreased reliance on analytic processes and, therefore, be more likely to conclude in favor of their first impressions.

To test this model, we adopted an information-order paradigm that has been used repeatedly in the medical education literature but with conflicting results. Cunningham et al.¹⁶ used a

procedure equivalent to our “generated” condition and discovered a primacy effect in physician judgments. In contrast, Chapman, Bergus, and Elstein¹⁷ used a procedure analogous to our “extreme” condition and discovered a recency effect in physician judgments. Our inclusion of 4 experimental conditions represented an attempt to better understand the reason for these conflicting results. We were able to replicate both effects, making it appear that procedural differences will influence the magnitude and direction of information-order effects. Other factors that are known to have an influence on information-order effects have been summarized by Hogarth and Einhorn.¹⁸ More important for the current article is that regardless of the procedure adopted, older physicians were more influenced by the information presented early in the case than were younger physicians, thereby suggesting that experience may lead to a decreased use of analytic processes. Overall accuracy cannot be determined with confidence as a result of

the methodologic design, but the literature cited above would suggest that too great a reliance on nonanalytic processing may yield an increase in errors attributable to premature closure.^{5,15}

There are limitations to this study:

- The sample is small. Still, the effect size is large enough that the study was sufficiently powered to reveal the hypothesized difference.
- The number of cases is low. Situation being a greater determinant of behavioral variability than people, we would have ideally used more cases. However, time was constrained, and we were hesitant to bias participants toward non-analytic processes by enforcing strict time limits on their responses.
- The 2 cohorts of physicians constitute a fairly narrow age range at the upper end of the experience scale. If anything, this should have lessened the probability of finding significant differences, but it was deemed acceptable because all longitudinal studies of aging suggest that age-related cognitive changes take place gradually rather than following a threshold model.¹⁹ Furthermore, the surprisingly large primacy effect in the most ecologically valid condition (the generated condition) leaves ample room for the existence of an even smaller primacy effect in younger cohorts. It is worth noting that in Cunningham et al.'s study, medical residents (ie, the youngest of qualified physicians) were the only group not to show a primacy effect.¹⁶

• The ecological validity of the procedure is questionable. Although certainly not a direct replication of the clinical environment, it must be remembered that this study was conceived to gain experimental control over a phenomenon observed in physicians through the course of their normal practices. Across a range of studies now, from low ecological validity and high control to high ecological validity and low control, the findings have been fairly consistent.

Therefore, we conclude that despite these limitations, the results of this study better position

Lessons for Practice

- With increased experience comes an increased tendency to conclude in favor of one's first impressions.
- The extent to which first impressions influence decision making is determined by the presence or absence of a diagnostic suggestion and one's confidence in the diagnosis.
- Educators should design continuing education activities that take into account experience-specific cognitive tendencies.

us to begin to question whether intentional adoption of particular diagnostic strategies can be taught and implemented to improve the clinical performance of struggling physicians. Viewing experience as a double-edged sword, as we propose, creates the opportunity to promote adoption of continuing education efforts more sophisticated than simply increasing the volume of information made available to experienced physicians. In general, should our model of expertise prove accurate, it would lead us to advocate recognizing the unique strengths experience provides while developing and investigating continuing education strategies that specifically support the controlled aspects of clinical diagnosis, thus counteracting the tendencies of individuals for whom medical practice has become excessively automated. This would create the potential for continuing education efforts and rehabilitation of failing physicians to be directed in a more efficacious manner.

Other work we have performed has revealed that medical trainees can be helped to overcome the biases that arise from too heavy a reliance on one form of processing over another.^{12,20} In multiple studies, diagnostic accuracy has been shown to

improve through the simple provision of instructions to list relevant features, for example, or through warning novice diagnosticians to trust a sense of familiarity in addition to carefully considering the diagnostic features presented (and providing them with practice in doing so). Whether repeated reinforcement of such simple manipulations (or analogous instruction) would be effective in more experienced physicians remains to be tested. Additional possibilities and implications that remain necessarily speculative at this point have been included in previous writings.^{5,6}

References

1. Kralj B. Physician human resources in Ontario: The crisis continues. *Ont Med Rev* 2001; 68:19–28.
2. Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: The relationship between clinical experience and quality of health care. *Ann Intern Med* 2005; 142:260–273.
3. Norman GR, Eva KW. Does clinical experience make up for failure to keep up to date? [editorial] *ACP J Club* 2005; 142:A8–A9.
4. Hobus PPM, Schmidt HG, Boshuizen HPA, Patel VL. Contextual factors in the activation of first diagnostic hypotheses: Expert-novice differences. *Med Educ* 1987; 21:471–476.
5. Eva KW. The aging physician: Changes in cognitive processing and their impact on medical practice. *Acad Med* 2002; 77(suppl): S1–S6.
6. Eva KW. Stemming the tide: Cognitive aging theories and their implications for continuing education in the health professions. *J Contin Educ Health Prof* 2003; 23:133–140.
7. Jacoby LL, Jennings JM, Hay JF. Dissociating automatic and consciously controlled processes: Implications for diagnosis and rehabilitation of memory deficits, In: Hertel D, Johnson MK, eds. Basic and applied memory research: Theory in context. vol. 1. Mahwah, NJ: Erlbaum, 1996:161–193.
8. Jacoby LL. Deceiving the elderly: Effects of accessibility bias in cued-recall performance. *Cogn Neuropsychol* 1999; 16:417–436.
9. Eva KW. What every teacher needs to know about clinical reasoning. *Med Educ* 2005; 39:98–106.
10. Norman GR, Brooks LR. The non-analytical basis of clinical reasoning. *Adv Health Sci Educ Theory Pract* 1997; 2:173–184.
11. Wigton RS. Social judgement theory and medical judgement. *ThinkReason* . 1996; 2:175–190.
12. Ark TK, Brooks LR, Eva KW. Giving learners the best of both worlds: Do clinical teachers need to guard against teaching pattern recognition to novices? *Acad Med* 2006; 81:405–409.
13. Norman GR, Brooks LR, Colle CL, Hatala RM. The benefit of diagnostic hypotheses in clinical reasoning: Experimental study of an instructional intervention for forward and backward reasoning. *Cogn Instruct* 2000; 17:433–448.
14. Norman G. Research in clinical reasoning: Past history and current trends. *Med Educ* 2005; 39:418–427.
15. Caulford PG, Lamb SB, Kaigas TB, Hanna E, Norman GR, Davis DA. Physician incompetence: Specific problems and predictors. *Acad Med* 1994; 69(suppl):S16–S18.
16. Cunningham JPW, Turnbull JM, Regehr G, Marriott M, Norman GR. The effect of presentation order in clinical decision making. *Acad Med* 1997; 72(suppl):S40–S42.
17. Chapman GB, Bergus GR, Elstein AS. Order of information affects clinical judgment. *J Behav Decis Making* 1996; 9: 201–211.
18. Hogarth RM, Einhorn HJ. Order effects of belief updating: The belief adjustment model. *Cogn Psychol* 1992; 24:1–55.
19. Schaie KW. Intellectual development in adulthood: The Seattle longitudinal study. New York: Cambridge University Press, 1996.
20. Eva KW. The influence of differentially processing information on diagnostic decision-making [doctoral dissertation]. Hamilton, Ont: McMaster University, 2001.