

Capsules and Comments

The Role of Decision Strategies in Understanding Professionals' Susceptibility to Judgment Biases

ANDREW D. CUCCIA* AND GARY A. MC GILL†

1. Introduction

The purpose of this paper is to examine whether professionals' ability to structure a judgment task affects the predictions of the belief-adjustment model (Hogarth and Einhorn [1992]). Although the model predicts that evidence evaluation order will significantly influence judgments, it implicitly assumes that evaluation order is a contextual factor incidental to the judgment task. However, the identification and organization of evidence are often important strategic components of accounting judgment tasks over which professionals exercise control. For example, tax professionals structure the legal research task by using databases that allow

*University of Oklahoma; †University of Florida. We thank A. Ahmed, S. Asare, M. Chatham, C. Copp, J. Demski, A. Gramling, S. Glover, K. Hackenbrack, J. Hunton, V. Johnson, R. Knechel, M. Lipe, W. Messier, M. Peecher, D. Prawitt, M. Prietula, C. Purdie, M. Roberts, I. Solomon, R. Tubbs, R. Worsham, two anonymous referees, and seminar participants at the University of Arkansas, Brigham Young University, the University of Florida, Oklahoma State University, and the University of Oklahoma for helpful comments and suggestions. G. McGill gratefully acknowledges the financial support provided by the Deloitte & Touche/University of Florida Partners-in-Research program and the Price Waterhouse Visiting Tax Professor program.

direct access to evidence used in supporting preferred reporting positions. We predict that allowing tax professionals control over evidence evaluation order will facilitate use of familiar and purposeful decision strategies and mitigate susceptibility to recency.¹

We report the results of two experiments. In the first, we manipulated tax professionals' ability to control the evidence evaluation process in a familiar tax judgment task, finding that the ability to control the order in which conflicting evidence was evaluated eliminated recency. In the second, we changed the judgment task to a relatively unfamiliar, non-professional context, while maintaining the underlying structure of the task. In this context, with subjects having no specialized knowledge or experience, allowing control over the evaluation process failed to mitigate recency. Taken together, these results suggest that recency was dependent on the interaction of control over the judgment task and the decision context. Recency was observed only when subjects had no context-relevant knowledge or experience or were precluded from structuring the task to facilitate its use.

2. *Background and Hypotheses*

The susceptibility of judgments to recency has been documented in a variety of accounting contexts, including assessments of internal controls (Ashton and Ashton [1988], Butt and Campbell [1989], and Trotman and Wright [1996]), the fair value of assets (Tubbs, Messier, and Knechel [1990], Messier [1992], Krull, Reckers, and Wong-On-Wing [1993], and Messier and Tubbs [1994]), firm viability (Asare [1992], Messier [1992], Kennedy [1993], Cushing and Ahlawat [1996], and Trotman and Wright [1996]), and tax reporting (Christian and Reneau [1990] and Pei, Reckers, and Wyndelts [1990; 1992]). Prior research has shown that this bias can be mitigated by environmental and task characteristics such as response mode (Tubbs, Messier, and Knechel [1990] and Trotman and Wright [1996]) and task complexity (Trotman and Wright [1996]), the interaction of these task characteristics with individual factors such as experience (Krull, Reckers, and Wong-On-Wing [1993], Messier and Tubbs [1994], and Trotman and Wright [1996]), or by encouraging additional effort (Kennedy [1993] and Cushing and Ahlawat [1996]). However, absent an external mitigating influence, the conclusion has been that accounting judgments generally are susceptible to recency.

¹We use the term "decision strategy" broadly to refer to the processes used to construct a solution to a problem. Decision strategies can include both mental and physical processes aimed at identifying, collecting, and evaluating evidence as well as documenting and justifying the preferred solution (Payne, Bettman, and Johnson [1988] and Payne et al. [1995]). Other terms have been used to refer to the same underlying construct. These include "procedural schema" (Newell and Simon [1972], Nisbett and Ross [1980], Ableson [1981], Taylor and Crocker [1981], and Glaser [1984]), "templates" and "decision guides" (Gibbins [1984]), and "knowledge structures" (Gibbins [1987]).

These studies test predictions of the belief-adjustment model (Hogarth and Einhorn [1992]). This model proposes that when people update beliefs, evidence is encoded relative to an initial anchor and the weight the evidence is given depends, in part, on whether it is consistent with prior beliefs (i.e., the anchor), and how strongly the prior belief is held. The weight given to evidence *contrary* to prior beliefs is greater the stronger the prior belief is held (i.e., the "contrast" effect). Thus, if a short series of conflicting evidence is reviewed sequentially, greater weight will be placed on evidence considered later because the evidence reviewed first will form the basis for prior beliefs, leading to a greater contrast between the anchor and the subsequent contrary evidence. The model implicitly assumes a passive decision maker in search of uncovering the truth, who is unaware of the nature and extent of existing evidence and has no control over the order in which evidence is evaluated. Recency is hypothesized to result from the contrast between the decision maker's prior belief and the surprisingly encountered new evidence, and the effort required to sequentially integrate the two.

Although these assumptions may be descriptive of some accounting tasks (e.g., detailed compliance testing and substantive tests of details), they may not be descriptive of many critical decision-making tasks applicable to professional accountants. For example, the accountant is often a goal-oriented rather than passive decision maker (Bonner and Pennington [1991] and Solomon and Shields [1995]), with an incentive to propose and defend favored propositions (Gibbins [1984; 1987] and Davis and Solomon [1989]). For example, a tax practitioner has a professional obligation to act as a client advocate and to interpret the law in a client's best interests (*AICPA* [1989]). Prior research suggests this role has an effect on actual judgments (Johnson [1993] and Cuccia, Hackenbrack, and Nelson [1995]). Further, decision makers often control the access to and evaluation of evidence. For example, several tax databases containing legislative and judicial guidance on which tax positions may be based are organized to allow direct access to selected pieces of evidence. If an accounting judgment is not characterized by a sequential anchoring and adjustment process involving seemingly randomly encountered evidence in a relatively passive search for truth, predictions of a model based on those assumptions may not be applicable to that judgment. Similarly, external validity is impaired when such a process is imposed on professional subjects or when the ability to use a preferred decision strategy is precluded in the laboratory.

Accounting professionals develop a repertoire of decision strategies consistent with their own goals as well as characteristics of the task and its environment (Gibbins [1984], Davis and Solomon [1989], Solomon and Shields [1995], Payne, Bettman, and Johnson [1988; 1992], Payne et al. [1992], and Payne et al. [1995]). As client advocates, tax professionals develop strategies to identify, organize, and present authoritative evidence (court cases, legislation, and administrative rulings) to build arguments that best justify client-favorable reporting positions (*AICPA*

[1989]). These strategies may vary because of differences in personal characteristics and prior experiences (Larrick, Morgan, and Nisbett [1990] and Kruglanski [1989]).²

Because an individual's ability to process information depends on the congruence of the information's organization with its intended use (Bettman, Payne, and Staelin [1986]), we expect evidence evaluation order is an important factor in tax professionals' ability to use their preferred decision strategies. Development of decision strategies in this context, and their dependence on order, are facilitated by the use of several legal databases that allow the tax practitioner to discern the nature and extent of available evidence and to control the order in which it is to be evaluated. Therefore, the identification and organization of evidence becomes an important component of the decision strategy rather than an incidental, situational factor as suggested by the belief-adjustment model. We expect that allowing control over evidence evaluation order will facilitate use of a preferred decision strategy, thereby reducing the complexity of the judgment task and susceptibility to recency. This leads to the following hypothesis.

H1: The ability to control evidence evaluation order in structuring a familiar professional task will mitigate recency in tax professionals' evaluations of conflicting evidence.

The effect predicted by *H1* is based on the ability of the decision maker to control the order in which evidence is evaluated. In the natural environment, this ability often is confounded with awareness of the nature and extent of the conflicting evidence. In the context of the belief-adjustment model, the mere expectation of encountering conflicting evidence can mitigate the contrast effect necessary to produce recency. Consequently, we explicitly examine whether the awareness of the information's existence alone, as opposed to the control over evaluation order such awareness facilitates, is sufficient to mitigate recency. This competing hypothesis is formally stated as follows:

H2: Pre-evaluation information regarding the existence and extent of conflicting evidence in a familiar professional task is sufficient to mitigate recency in tax professionals' evaluation of conflicting evidence.

We do not expect that control over evaluation order would mitigate recency in less familiar contexts where professionals do not possess task-

² Some tax professionals may prefer to examine first cases decided in favor of the taxpayer in order to structure an argument around the critical issues in those cases, emphasizing similarities between them and the issue at hand. Others may prefer to examine first cases decided against the taxpayer, with the goal of distinguishing those cases from the issue at hand, or to quickly determine if any of those cases clearly precludes reaching the preferred conclusion.

specific knowledge or experience. If task-structuring ability does not mitigate recency in an unfamiliar context, we can conclude that the effect of control over evaluation order is dependent on a particular decision context, i.e., the ability to apply a familiar, context-specific decision strategy. This leads to the following hypotheses:

- H3*: The ability to control evidence evaluation order in structuring an unfamiliar task will fail to mitigate recency in tax professionals' evaluations of conflicting evidence.
- H4*: Pre-evaluation information regarding the existence and extent of conflicting evidence in an unfamiliar task is sufficient to mitigate recency in tax professionals' evaluations of conflicting evidence.

We performed two experiments to test these hypotheses. The first involved a relatively familiar tax judgment task in which we expected recency only when control over evidence order was precluded. The second experiment involved a similarly structured judgment task in a relatively unfamiliar, nonprofessional context in which we expected subjects to exhibit recency regardless of control over evidence order.

3. *Experiment 1*

We tested *H1* and *H2* by examining tax professionals' judgments regarding the treatment of a transaction with ambiguous tax consequences. We administered the experimental task via computer using an interactive software program described by Brucks [1988] that was tailored for this study. Subjects were asked to assess the taxability of a personal injury settlement received by a hypothetical client and to make a related reporting recommendation. The settlement was the result of defamatory public statements aimed at the client by investors in an unsuccessful real estate development project. Although Internal Revenue Code Section 104 excludes from taxable income certain amounts received on account of personal injury, its applicability to the settlement was ambiguous at the time of the experiment.

Subjects read a brief description of the client, the facts leading to the settlement, the terms of the settlement, and the statutory language of Section 104 before providing preliminary estimates of the probability the settlement would be found excludable from income if litigated. Next, the computer program presented subjects with four court cases related to the taxation of amounts received on account of personal injury. The cases averaged two computer screens—about 235 words in length. The facts in the four cases (all based on actual court cases) were similar to those in the scenario, with two cases supporting exclusion treatment and two cases disallowing exclusion treatment.³ After reviewing the four

³ The size of order effects can be affected by the relative diagnostic strength and saliency of the evidence (Hogarth and Einhorn [1992]). Thirty-eight managers from the same firm

cases, subjects provided a final probability estimate and made a reporting recommendation. Last, subjects provided certain attitudinal and demographic information. The experiment was administered by mailing the instrument to subjects' offices along with a booklet explaining the purpose of the study and instructions regarding use of the program. The average time to complete the entire task was about 30 minutes.

We manipulated three variables between subjects producing a 2 (evidence order) \times 2 (ability to control evidence evaluation order in structuring the task) \times 2 (availability of pre-evaluation information) fully crossed experimental design. Approximately half the subjects reviewed cases in which a settlement was found taxable (excludable) before those cases finding the settlement excludable (taxable). Note that evidence order was only partially manipulated; subjects given the ability to structure the task self-selected into a specific evidence order condition.

We manipulated the ability to structure the task across subjects by allowing half the subjects to select the order in which they assessed the judicial evidence (e.g., *PPNN* or *NNPP*). Half the subjects viewed the evidence in a predetermined order.⁴ These options follow those available in practice. For example, computer-based legal databases provide users a choice of examining only cases that (1) follow the case in question, (2) do not follow the case in question, or (3) all cases. To the extent that the available options did not include the order preferred by some subjects, this was a weak manipulation and our results are biased against finding support for *H1*.

We manipulated the availability of pre-evaluation information across subjects by providing half the subjects with short descriptions of all four cases, and the court's decision in each, before the cases were evaluated in detail. All four descriptions were presented on a single computer screen in a format similar to that used in the annotations of the major legal databases. The individual case descriptions appeared on the screen in an order randomized across subjects. The remaining subjects received no information regarding the nature or extent of relevant judicial evidence before sequentially assessing the actual cases. Subjects with pre-

as our experimental sample independently rated the relevance of each case to the hypothetical scenario on a scale of 1 (not important or relevant) to 7 (definitely relevant or "on point"), with the order in which the cases were rated counterbalanced across raters. Mean ratings ranged from 5.03 to 5.34, with no significant differences across cases ($p = .64$) nor a significant interaction between case and evaluation ($p = .98$).

⁴Two observations regarding the task manipulation should be noted. First, although subjects given the ability to structure the evaluation task self-selected into a specific evidence order condition, all subjects were systematically assigned to task conditions allowing or prohibiting control over the order evaluation process. Second, subjects prohibited from controlling evidence evaluation order may have been assigned to an order condition they would have chosen if allowed. Therefore, systematically altering subjects' control over the order evaluation process is a weak manipulation of their ability to evaluate the evidence in their preferred order, biasing the results against support for *H1*.

evaluation information were given control over evidence order by asking them which case they wanted to review first. Subjects without pre-evaluation information were asked if they wanted the computer to search first for cases in which an award was found to be excludable from income or for cases in which an award was found taxable. Sample case materials appear in Appendix A.

We collected two additional control variables: subjects' initial probability estimates that a court would find the receipt in question excludable from gross income and a general measure of client advocacy. Because the belief-adjustment model addresses the influence of new evidence on decision makers' beliefs, or changes in beliefs, we controlled for subjects' initial beliefs by asking subjects to provide judgments after they reviewed the case scenario and the actual, inconclusive statutory language of Section 104 but before they reviewed any judicial evidence. Because tax professionals' interpretations of authority have been shown to be sensitive to the importance they attach to their role as client advocate (e.g., Johnson [1993] and Cuccia [1994; 1995]), we controlled for subjects' client advocacy attitudes. We measured client advocacy using subjects' responses to an 18-item scale (Johnson [1993]), collected after all evidence was evaluated and judgments made. These responses were factor analyzed with factor scores used as the measure of client advocacy.⁵

We collected two dependent variable measures after subjects reviewed all four cases: a probability estimate that a court would find the settlement excludable from taxable income and a reporting recommendation to (1) include the receipt in gross income, (2) exclude, but disclose, the receipt, or (3) exclude the receipt without disclosure. The presentation order of these three options was randomized across subjects.

3.1 PRELIMINARY ANALYSES

Ninety-four tax professionals from several offices of one Big Five *CPA* firm participated in experiment 1. On average, these subjects were 30 years of age with 5 years of tax experience. They were somewhat unfamiliar (mean self-rating = 2.6; 1 = not familiar and 7 = very familiar) with the tax treatment of personal injury settlements, alleviating concerns that they based their experimental judgments on knowledge brought by them to the study rather than on the evidence provided. Subjects rated the tax research task and context as realistic (mean response = 5.4; 1 = very contrived, 7 = very realistic). The mean initial estimate of the probability a court would find the hypothetical settlement excludable from income

⁵ Items in the scale generally capture subjects' beliefs regarding appropriate behavior when the tax law and standards are ambiguous. Subjects' responses to the 18-item scale were found to load highly on a single factor that explained 25% of the total variance in responses. Cronbach's [1951] alpha coefficient of 0.80 was similar to the 0.82 reported by Johnson [1993] and satisfies Nunnally's [1994] criterion for construct reliability.

was 48%. The mean final estimate of this same probability was 57%, with 39% of the subjects recommending exclusion with disclosure, 34% recommending exclusion without disclosure, and 27% recommending inclusion. These results suggest that the subjects viewed the tax treatment of the settlement as ambiguous. General tax experience, familiarity with the taxation of damages, ratings of task realism, preliminary assessments of exclusion probability, age, gender, and client advocacy scores did not differ across treatments at traditional significance levels, suggesting that we achieved random assignment of subjects across treatments.

3.2 HYPOTHESES TESTS

We examined *H1* and *H2* by performing an *ANCOVA* test for the effects of evidence order, structuring ability, and pre-evaluation information on subjects' final estimates of the probability that a court would find the hypothetical settlement excludable from taxable income. We included subjects' initial probability estimates and client advocacy measures in the analysis as covariates. Results are presented in table 1, panel A.⁶

The overall model is significant ($p < .001$) in predicting subjects' final probability estimates. As predicted by *H1*, the effect of evidence order is moderated by subjects' ability to control evidence order ($p = .063$).⁷ Simple effects tests presented in panel B of table 1 show that the effect of evidence order on final probability estimates was greater for subjects presented the judicial evidence in a predetermined order (51 versus 62, $p = .089$) than for subjects able to choose the order in which evidence was examined (58 versus 53, $p = .765$). Finally, a planned contrast of the cell means supports ($p = .066$; not reported in table 1) the trend suggested in *H1* (i.e., the highest estimates by those forced to evaluate the evidence in the *NNPP* order, followed by those with the ability to control the evidence order, with those forced to evaluate the evidence in the *PPNN* order having the lowest estimates).

An alternative explanation of the recency-mitigating effect of structuring ability is that when subjects were allowed to choose the evidence evaluation order, stronger client advocates self-selected into the *PPNN* order because of confirmation proneness, i.e., the tendency to provide spurious support for a preferred hypothesis by selectively searching or assim-

⁶Alternatively, we tested the hypotheses by examining subjects' belief revisions, or the difference between initial and final probability estimates (e.g., Messier and Tubbs [1994]). The reliability of difference scores can be less than that of their individual components, thereby reducing the power of tests employing them. Further, difference scores can raise statistical and construct validity concerns by confounding the effects of the individual components, thereby losing information conveyed by each (Bedeian et al. [1994]). Our results are not qualitatively different using difference scores.

⁷We report *p*-values based on one-tailed tests in our examination of main and interaction effects involving evidence order (other than two-way interactions for which we have no predictions) because the hypotheses involve directional predictions.

TABLE 1
Effects of Evidence Order, Structuring Ability, and Pre-Evaluation Information on Final Judgments in a Tax Judgment Task: Experiment 1^a

Panel A: Analysis of Covariance		
Source of Variance	<i>F</i>	<i>p</i> > <i>F</i>
Evidence Order ^b	0.38	.269*
Structuring Ability ^c	0.02	.883
Pre-Evaluation Information ^d	0.24	.628
Order × Structuring Ability	2.40	.063*
Order × Pre-Evaluation Information	0.62	.783*
Structuring Ability × Pre-Evaluation Information	1.29	.259
Order × Structuring Ability × Pre-Evaluation Information	0.52	.472
Initial Estimate ^e	29.72	.000*
Client Advocacy ^f	5.34	.023*
Overall Model	4.07	0.000
<i>R</i> ²		.30
Number of Observations		94

Panel B: Simple Effects Tests and Experiment Means^a

	Evidence Order ^b		Statistical Significance of Difference*
	<i>PPNN</i>	<i>NNPP</i>	
Structuring Ability ^c			
Not Allowed	51	62	<i>F</i> = 1.88; <i>p</i> > <i>F</i> = .089
Allowed	58	53	<i>F</i> = 0.53; <i>p</i> > <i>F</i> = .765

*Reported *p*-value is one-tailed because test corresponds to a directional hypothesis.

^aThe response variable is the final probability estimate that a court would find the receipt in question excludable from gross income, where 0 = "no chance of being found excludable" and 100 = "would definitely be found excludable."

^bSubjects evaluated positive evidence prior to (*PPNN*) or subsequent to (*NNPP*) negative evidence.

^cSubjects were or were not able to choose the order in which to evaluate evidence.

^dSubjects did or did not receive information regarding the nature and extent of existing evidence before evaluating the evidence in detail.

^eInitial estimate refers to subjects' probability estimates prior to the review of any evidence.

^fPrincipal component factor score extracted from an 18-item client advocacy scale (Johnson [1993]). Higher scores represent a higher level of client advocacy.

ilating information (Arkes [1991]).⁸ Thus, the increased weight placed on negative evidence due to recency may have been offset by the weight on positive evidence due to confirmation proneness. We tested this competing explanation by examining the relation between client advocacy and the evidence order selected by subjects. Although 57% of those

⁸ See Klayman and Ha [1987], Church [1990], and Klayman [1994] for reviews of the confirmatory processes literature. Johnson [1993] and Cloyd and Spilker [1999] provide evidence that tax professionals' client advocacy role may influence their evaluations of authoritative evidence in a manner consistent with confirmation proneness. If client advocacy leads to confirmation proneness in tax professionals, it may be manifested not only in the *evaluation* of available evidence but in the *order* in which evidence is evaluated. Our purpose was not to test for confirmation proneness but simply to rule out this alternative explanation of our results.

