The Dirt on Coming Clean: Perverse Effects of Disclosing Conflicts of Interest

Daylian M. Cain, George Loewenstein, and Don A. Moore

ABSTRACT

Conflicts of interest can lead experts to give biased and corrupt advice. Although disclosure is often proposed as a potential solution to these problems, we show that it can have perverse effects. First, people generally do not discount advice from biased advisors as much as they should, even when advisors' conflicts of interest are disclosed. Second, disclosure can increase the bias in advice because it leads advisors to feel morally licensed and strategically encouraged to exaggerate their advice even further. As a result, disclosure may fail to solve the problems created by conflicts of interest and may sometimes even make matters worse.

1. INTRODUCTION

Conflicts of interest occur when individuals' professional responsibilities diverge from their personal interests (or when different professional responsibilities clash). Attorneys often face conflicts of interest when they advise clients on whether to pursue legal action. Doctors face conflicts of interest when they advise patients on whether to get procedures that they will profit from performing. Stock analysts face conflicts of interest when they are in a position to benefit financially from promoting a stock on which they are supposed to provide an impartial evaluation. Ac-

DAYLIAN M. CAIN is Doctoral Candidate of Organizational Behavior and Theory at Tepper School of Business, Carnegie Mellon University. GEORGE LOEWENSTEIN is Professor of Psychology and Economics at the Department of Social Decision Sciences, Carnegie Mellon University. DON A. MOORE is Associate Professor of Organizational Behavior and Theory at Tepper School of Business, Carnegie Mellon University. The authors wish to thank Max Bazerman, Lloyd Tanlu, Roland Benabou, and Ben Haimowitz for helpful comments, the American Accounting Association for financial support, and Sam Swift for superb assistance in running the experiment.

[*Journal of Legal Studies*, vol. 34 (January 2005)]
 0047-2530/2005/3401-0001$01.50

counting firms and their employees face conflicts of interest when they audit the same companies to which they provide consulting services. These specific conflicts of interest, and many others, have received substantial scrutiny as a result of recent historical developments such as rising medical costs, scandals involving stock analysts, and the highly publicized collapse of firms such as Enron and WorldCom in which conflicts of interest were perceived to have played a role.

Diverse solutions have been proposed to address the problems caused by conflicts of interest (see Issacharoff, forthcoming; Coffee 2004; Davis and Stark 2001; Stark 2000). For example, the Sarbanes-Oxley Act (U.S. Congress 2002) limits the ability of accounting firms to provide both auditing and consulting services to the same client. Similarly, rules for financial service firms have been proposed that would limit compensation that analysts receive from investment banking activity and restrict analysts from trading stocks that they cover. Most of the responses that have been proposed for the problems caused by conflicts of interest, however, incorporate a common element: they include disclosure as a critical ingredient.

Medical journals, for example, require researchers to disclose the sources of their research funding. Financial media sources such as CNBC and CNNfn now require stock analysts to disclose their conflicts of interest when they offer televised advice on stocks. And an entire section of Sarbanes-Oxley (Title IV) is dedicated to enhanced disclosure by corporations and their auditors. Finally, one key feature of most campaign finance reform legislation, including the McCain-Feingold Act (U.S. Congress 2001), is to mandate public disclosure of political contributions.

Common sense suggests that recipients of advice will benefit from being more fully informed when they are made aware of an advisor's conflict of interest. According to its most staunch supporters, disclosure reduces the need for other remedies and even eliminates the need for any conflict-of-interest regulation whatsoever. In the words of former U.S. senator Philip Hart, disclosure works by "revealing the possibility of . . . conflict, leaving it to the voter to decide whether the conflict has influenced the official acts of the congressman or senator" and to the congressman or senator to decide how to respond (Hart 1975, p. 2019).

1.1. Disclosing Conflicts of Interest

Disclosure is a popular response to conflicts of interest in part because it promises something to everyone. For recipients of advice, disclosure

provides potentially useful information. Healy and Palepu (2001, p. 412), for example, note that "regulators may be concerned about the welfare of financially unsophisticated investors. By creating minimum disclosure requirements, regulators reduce the information gap between informed and uninformed."[1] It stands to reason that knowledge of a conflict of interest should permit recipients of biased advice to discount that advice and make better subsequent decisions.

The professionals who provide information and advice are also likely to embrace disclosure as the lesser of evils since it generally involves minimal disruption of the status quo. Physicians will prefer disclosing gifts from pharmaceutical companies (or disclosing payments for referring patients to clinical trials) to actually eschewing such benefits. Likewise, stock analysts will prefer to disclose industry contacts rather than eliminate them. And auditors almost certainly would have preferred to disclose any consulting services they provided to their auditing clients rather than surrender this lucrative source of revenue or split themselves into independent auditing and consulting units.

Disclosure offers a further benefit to both advisors and to policy makers: it diminishes both parties' responsibility for adverse outcomes.[2] If patients agree to participate in clinical trials from which they know their physicians benefit, if investors rely on the advice of analysts whose ties to industry have been disclosed, or if investors continue to accept audit reports from companies they know have conflicts of interest, it could be argued that these patients and investors should be held responsible for any negative consequences that result; caveat emptor.

Perhaps, however, the benefits of disclosure should not be accepted quite so quickly. For disclosure to be effective, the recipient of advice must understand how the conflict of interest has influenced the advisor and must be able to correct for that biasing influence. In many important situations, however, this understanding and ability may be woefully lacking. For example, imagine a patient whose physician advises, "Your life is in danger unless you take medication X," but who also discloses, "The

1. Analytical research on disclosure is discussed by Verrecchia (2001) and Dye (2001). Empirical research on diverse forms of disclosure, not just that involving conflicts of interest, is reviewed by Healy and Palepu (2001).

2. Likewise, "Consumer advocates hailed [the requiring of warning labels on cigarette packages] as a great victory. But since the labels first appeared, the industry has fended off smokers' suits by citing [the warning labels] as evidence that smokers should have known the risks. What was intended as a burden on tobacco became a shield instead" (Action on Smoking and Health 1997).

medication's manufacturer sponsors my research." Should the patient take the medication? If not, what other medication? How much should the patient be willing to pay to obtain a second opinion? How should the two opinions be weighed against each other? The typical patient may be hard-pressed to answer such questions.

And what is the impact of disclosure on providers of advice? In the example just given, is it possible that the physician's behavior might be affected by disclosure? For example, might the physician be more likely to exaggerate the danger of not taking the medication in order to neutralize the anticipated "warning" effect of the disclosure? Such exaggeration could be more likely, even if the physician had only the best of intentions in persuading her patient to take the medication. And might the physician feel less personally responsible for promoting the patient's interest once the patient has been warned?

In this paper, we enumerate psychological factors that may undermine disclosure's effectiveness as a remedy for the problems caused by conflicts of interest. These factors may even cause disclosure to backfire, harming rather than helping the recipients of advice. We document such a perverse effect in an experiment designed to replicate the major features of many situations characterized by conflicts of interest.

1.2. Potential Pitfalls of Disclosure

Crawford and Sobel (1982) present a theoretical analysis of a situation that could be characterized as a fully disclosed conflict of interest. In their model, an agent (who we will call the "estimator") attempts to estimate the value of an uncertain quantity and is then rewarded on the basis of the accuracy of her estimate. The estimator is provided with information by a second individual (the "advisor") who, however, has incentives that are mutually understood to be different from those of the estimator.[3] Crawford and Sobel show that the estimator's use of the information provided by the advisor and the expected utility to each agent increase as a function of the alignment of the two agents' interests. Probably because it would require a myriad of extra assumptions and would seriously complicate their already complicated analysis, Crawford and Sobel do not examine the case of asymmetric information in which the estimator does not have full information about the advisor's incen-

3. Crawford and Sobel refer to the two agents as the "receiver" and "sender." We use the terms "estimator" and "advisor" to draw a connection between their work and our experimental study, in which our terms more accurately describe the experimental setup.

tives. But understanding what happens in this situation and examining the effects of moving from this situation to one in which the conflict of interest is common knowledge is key to understanding the effects of disclosure.

What should one expect to happen when conflicts of interest are disclosed? Revelation of the fact that interests are not aligned should logically lead estimators to be more suspicious of their advisors and to place less weight on the information that advisors provide. If advisors with conflicts of interest do indeed provide biased advice, then a decrease in estimators' reliance on the biased information should tend to decrease advisors' payoffs and increase estimators' payoffs.

Such an analysis, however, makes a number of simplifying assumptions. Most important, it assumes that estimators know what to do with the information that is disclosed. There are grounds for skepticism regarding this assumption.

First, estimating the impact of a conflict of interest on an advice giver is an extraordinarily difficult problem that requires both economic and psychological insight. To properly estimate the degree to which a particular advisor is biased by a conflict of interest, one would want to know the extent to which the advisor embraces professional norms or is instead corrupt. One would also want to know how tempting the advisor finds the incentives for providing biased advice, and one would want to have an accurate mental model of how such incentives can bias advice. However, prior research suggests that most people have an incorrect understanding of the psychological mechanisms that transform conflicts of interest into biased advice. While most people think conflicts of interest are a problem of overt corruption, that is, that professionals consciously and intentionally misrepresent the advice they give so as to secure personal gain, considerable research suggests that bias is more frequently the result of motivational processes that are unintentional and unconscious (for summaries of this research, see Dana and Loewenstein 2003; Moore and Loewenstein 2004). Failing to appreciate the role of unconscious bias will cause estimators to underestimate the extent to which advice might be distorted.

Second, there is at least suggestive evidence that people tend to be naturally trusting and credulous toward their own advisors. In the domain of medicine, for example, research shows that while many people are ready to acknowledge that doctors might generally be affected by conflicts of interest, few can imagine that their own doctors would be affected (Gibbons et al. 1998). Indeed, it is even possible that disclosure

could sometimes increase rather than decrease trust, especially if the person with the conflict of interest is the one who issues the disclosure. Research suggests that when managers offer negative financial disclosures about future earnings, they are regarded as more credible agents, at least in the short term (Lee, Peterson, and Tiedens 2004; Mercer, forthcoming). Thus, if a doctor tells a patient that her research is funded by the manufacturer of the medication that she is prescribing, the patient might then think (perhaps rightly) that the doctor is going out of her way to be open or that she is "deeply involved" and thus knowledgeable. Thus, disclosure could cause the estimator to place more rather than less weight on the advisor's advice.

Third, even when estimators realize that they should make some adjustment for the conflict of interest that is disclosed, such adjustments are likely to be insufficient. As a rule, people have trouble unlearning, ignoring, or suppressing the use of knowledge (such as biased advice) even if they are aware that it is inaccurate (Wilson and Brekke 1994). Research on anchoring, for example, shows that quantitative judgments are often drawn toward numbers (the anchors) that happen to be mentally available. This effect holds even when those anchors are known to be irrelevant (Strack and Mussweiler 1997; Tversky and Kahneman 1974), unreliable (Loftus 1979), or even manipulative (Galinsky and Mussweiler 2001; Hastie, Schkade, and Payne 1999). Research on the "curse of knowledge" (Camerer, Loewenstein, and Weber 1989) shows that people's judgments are influenced even by information they know they should ignore. And research on what has been called the "failure of evidentiary discreditation" shows that when the evidence on which beliefs were revised is totally discredited, those beliefs do not revert to their original states but show a persistent effect of the discredited evidence (Skurnik, Moskowitz, and Johnson 2002; Ross, Lepper, and Hubbard 1975). Furthermore, attempts to willfully suppress undesired thoughts can lead to ironic rebound effects, in some cases even increasing the spontaneous use of undesired knowledge (Wegner 1994).

In sum, diverse lines of research suggest that estimators may not discount advice from biased advisors as much as they should when conflicts of interest are disclosed and that, in some circumstances, disclosure may even lead estimators to put greater weight on biased advice.

Turning to the advisors, there are two ways in which disclosure could potentially worsen the advice that they provide to estimators. The first involves the advisors' strategic response to the disclosure of their conflict of interest. Logically, it is not clear how self-interested advisors should

respond to disclosure of their conflict of interest. On one hand, disclosure might deter advisors from giving biased advice by increasing their concern that estimators (now thought to be alerted by disclosure) will completely discount extreme advice or attribute corrupt motives to advice that seems even remotely questionable. On the other hand, advisors might be tempted to provide even more biased advice, exaggerating their advice in order to counteract the diminished weight that they expect estimators to place on it; this strategic exaggeration is like expecting disclosure to cause one's audience to cover its ears and thus compensating for this by yelling even louder.

The second way in which disclosure could influence the behavior of advisors involves what we call "moral licensing."[4] While Crawford and Sobel's (1982) model assumes that both agents' motivations are purely selfish, real professionals often are also motivated by a desire to live up to the norms of their profession. Only a hardened cynic would believe that doctors are not motivated by concerns for their patients, that attorneys are indifferent regarding ethical and professional conduct, and that auditors care only about lining their pockets and not at all about fulfilling their legal obligation to provide unbiased audits. Indeed, the whole notion of a conflict of interest assumes that professionals do experience such a conflict between self-interest and professional responsibilities. To the degree that people care about their professional responsibilities, disclosing conflicts of interest can potentially backfire by reducing advisors' feelings of guilt about misleading estimators and thereby giving advisors moral license to bias advice even further than they would without disclosure. With disclosure of a conflict of interest, giving biased advice might seem like fair play. While most professionals might care about their clients, disclosure regulation can encourage these professionals to exhibit this concern in a merely perfunctory way, by disclosing instead of eliminating conflicts of interest.

In sum, there are good reasons to worry that disclosure might not mitigate the problems caused by conflicts of interest and might even exacerbate them, increasing the bias in advice offered without producing commensurate discounting on the part of estimators.

In addition, both economic and psychological factors should lead to

4. Monin and Miller (2001) discuss a concept similar to moral licensing that they call "self-licensing." They show that once people demonstrate that they are not morally corrupt in some way, they are more likely to display exactly this corruption on subsequent tasks. For example, when people are given the opportunity to exhibit egalitarianism, they will subsequently be more likely to act on racist or sexist stereotypes.

an increase in the variance of estimators' estimates as a result of disclosure. From an economic perspective, as Crawford and Sobel's (1982) analysis shows, estimators should discount advice to the extent that they believe the incentives of the advice giver diverge from their own incentives. Disclosing conflicts of interest should therefore cause estimators to fall back on their own private information. To the extent that this private information is unreliable, as is the case in our experiment, such estimates are likely to be more highly dispersed. From a psychological perspective, different advisors are likely to differ in the degree to which they are altruistic or self-interested, and different estimators are likely to differ in their beliefs about the extent of altruism or self-interest among advisors. Both forms of heterogeneity should, again, increase the variance of estimators' estimates when the conflict of interest is disclosed.

2. THE STUDY

2.1. Predictions

To investigate the impact of disclosure of conflicts of interest, we conducted an experiment in which subjects played one of two roles: estimator or advisor. Estimators attempted to estimate an uncertain quantity and were rewarded for accuracy. Advisors were provided with more information than estimators and were instructed to provide estimators with advice. In a control treatment, advisors, like estimators, were paid more when estimators answered accurately. This alignment of incentives was disclosed. In two conflict-of-interest treatments, advisors were paid more when the estimator responded with a high (relative to actual value) rather than an accurate estimate. We examined the impact of disclosure by disclosing this conflict of interest in one of the conflict-of-interest treatments but not in the other.

We test the following three predictions:

1. Estimators' estimates will be less reliant on advisors' advice with disclosure than without disclosure.
2. Advisors with conflicts of interest will give more biased advice under conditions with disclosure than without disclosure.
3. Estimators will make higher and more dispersed, and therefore less accurate, estimates with disclosure of conflicts of interest than without their disclosure, which will lead to (*a*) lower payoffs for estimators and (*b*) higher payoffs for advisors.

The first prediction describes rational behavior on the part of estimators and is consistent with a standard economic analysis. The second follows from the strategic and moral licensing mechanisms discussed in the previous section. The third derives from the reasons, also discussed in the previous section, that estimators are unlikely to adjust adequately for knowledge of the conflicting incentives of advisors when these are disclosed. In addition to these basic predictions, we also examine other factors, such as the effect of feedback on estimator and advisor payoffs.

2.2. Experimental Method

Participants were 147 undergraduate students at Carnegie Mellon University, recruited for pay ($6–$15 per hour, with an average of $10). They participated six to 10 at a time and were randomly assigned to the role of either advisor or estimator, which they retained throughout the experiment. The estimation task involved estimating the values of jars of coins. Estimators were paid according to the accuracy of their estimates, and advisors were paid, depending on the experimental condition, on the basis of either how accurate or how high (relative to actual values) the estimators' estimates were.

Participants were seated at cubicles and were given codes that kept their identities anonymous. There were six jars and thus six rounds, and the presentation order of the jars varied session by session. In each round, advisors took turns at closely examining a jar of coins and then completed an advisor's report. Each advisor's report contained the advisor's suggestion of the value of the jar in question and provided a space in which the estimator would respond with an estimate of the jar's worth. Once advisors wrote their suggestions on the advisor's reports, the reports were then handed to the experimenter, who shuffled them and gave one to each estimator. Each estimator got one advisor's report per round, with an equal probability of getting any advisor's advice, including that of the same advisor from whom they had received advice in the prior round. If there was an odd number of participants in a session, we had one more estimator than we had advisors, and one advisor's report was randomly selected for duplication.[5] These procedures were made transparent to participants before the experiment began.

After seeing the reports, the estimators saw the jar in question—but

5. Our analysis does not include these duplicates when examining what advisors did but does include duplicated advice when examining what estimators did (since a unique estimator's reaction to even duplicated advice is informative).

only from a distance of about 3 feet and only for about 10 seconds: the experimenter held the jar in front of estimators, turning the jar while walking along a line across the room and back. Estimators then attempted to estimate the value of the coins in the jar.

The amount of money in each of the six jars (M, N, P, R, S, and T) was determined somewhat arbitrarily to lie between $10 and $30, and advisors were informed of this range. Estimators were told that advisors had information about the range of actual values but were not given this range of values themselves. In fact, the values of the jars were M = $10.01, N = $19.83, P = $15.58, R = $27.06, S = $24.00, and T = $12.50. In the first three rounds, neither estimators nor advisors received feedback about their actual payoffs or about actual jar values. In each of the last three rounds, however, after advisors had given their advice and estimators had made their estimates, each advisor was shown the estimate of the estimator to whom their advice was given on the previous jar and, for each of the feedback rounds, the actual value of the jar in question was announced to everyone at the end of the round. Since payoff schedules (described below) were provided to all participants at the beginning of the experiment, feedback allowed both advisors and estimators to calculate how much money they had made in the previous round before continuing on to the next round. While estimators did not see the advisor's instructions, advisors saw a copy of the estimator's instructions and thus could also use feedback to calculate their estimator's payoffs.

Both estimators and advisors were paid on the basis of the estimator's estimates. Estimators were always paid on the basis of the accuracy of their own estimates. Advisors' remuneration depended on the condition to which they were assigned, as described in Tables 1 and 2. In the "accurate" condition, each estimator was paid according to how accurate the estimator's estimate was, and this was disclosed prominently on the advisor's report immediately under the advisor's suggestion. ("Note: The advisor is paid based on how accurate the estimator is in estimating the worth of the jar of coins.") In the "high/undisclosed" and "high/disclosed" conditions, each advisor was paid on the basis of how high the estimator's estimate was. This conflict of interest was not disclosed in the high/undisclosed condition but was prominently disclosed in the high/disclosed condition, immediately under the advisor's suggestion. ("Note: The advisor is paid based on how high the estimator is in estimating the worth of the jar of coins.") In addition to being remunerated on the basis of their estimators' estimates, all advisors had

Table 1. Payoff Function for Advisors in Accurate Condition and for All Estimators

Range of Estimator's Estimate from True Value ($)	Payoff ($)
.00–.50	5.00
.51–1.00	4.50
1.01–1.50	4.00
1.51–2.00	3.50
2.01–2.50	3.00
2.51–3.00	2.50
3.01–3.50	2.00
3.51–4.00	1.50
4.01–4.50	1.00
4.51–5.00	.50

Table 2. Advisors' Payoff Function in Conflict-of-Interest Conditions

Range of Estimator's Estimate above True Value ($)	Payoff ($)
.50–1.00	1.00
1.01–1.50	1.90
1.51–2.00	2.70
2.01–2.50	3.40
2.51–3.00	4.00
3.01–3.50	4.50
3.51–4.00	4.90
4.01–4.50	5.20
4.51–5.00	5.40
5.01+	5.50

an additional opportunity to earn money: after they had completed the report for each jar, advisors were asked to give their own personal best estimates of the true value of the coins in the jar and were rewarded on the basis of accuracy (see Table 3).

Participants were instructed that at the end of the experiment, one of the six rounds would be randomly selected to serve as the "payoff" round. Any money earned in that round (including earnings based on estimators' estimates, and for advisors, based in addition on their personal estimates) would be paid in cash in addition to a $7.50 base payment. Participants were encouraged to ask questions if they did not understand any of the instructions. Simple "yes" or "no" answers sufficed to answer the few questions that arose.

Table 3. Advisors' Payoff Function for Personal Estimate

Range of Advisor's Estimate from True Value ($)	Bonus Payment ($)
.00–.50	2.00
.51–1.00	1.75
1.01–1.50	1.50
1.51–2.00	1.25
2.01–2.50	1.00
2.51–3.00	.75
3.01–3.50	.50
3.51–4.00	.25

3. RESULTS

Results from the experiment were analyzed with repeated measures analyses of variance (ANOVAs) in which independent variables were (1) the experimental condition, which was manipulated between subjects, and (2) round (1–6), which was measured within subjects. We compared the accurate condition to the two conflict-of-interest conditions and compared the high/undisclosed condition to the high/disclosed condition.[6]

3.1. Advisors' Suggestions and Personal Estimates

As evident in Figure 1, advisors' suggestions differed substantially across the three conditions ($F(2, 59) = 9.76$, $p < .001$). The mean actual jar value (across the six jars) was $18.16, but the mean value of advice given by incentive condition across jars was $16.48 in the accurate condition, $20.16 in the high/undisclosed condition, and $24.16 in the high/disclosed condition.

As the first of these suggestions implies, there was a general tendency to underestimate jar values when incentives were aligned. Table 4 presents actual values for the six jars and compares these to the mean personal estimates of advisors in the accurate condition. Advisors tended to underestimate the value of the jars, in some cases quite dramatically. Below we address the problems that resulted from this underestimation.

Table 5 lists advisors' personal estimates and suggestions and com-

6. Our analyses omitted two suggestions that were extreme outliers ($.01 and $4,000). Both suggestions were from the same advisor who was in the high/disclosed condition. The $.01 suggestion is more than 3 standard deviations below the mean suggestion for that condition, and $4,000 is more than 500 standard deviations above the mean. All other suggestions, across all jars and all conditions, ranged between $2.75 and $96.00. There were other missing data points, and these caused degrees of freedom to fluctuate slightly between tests.

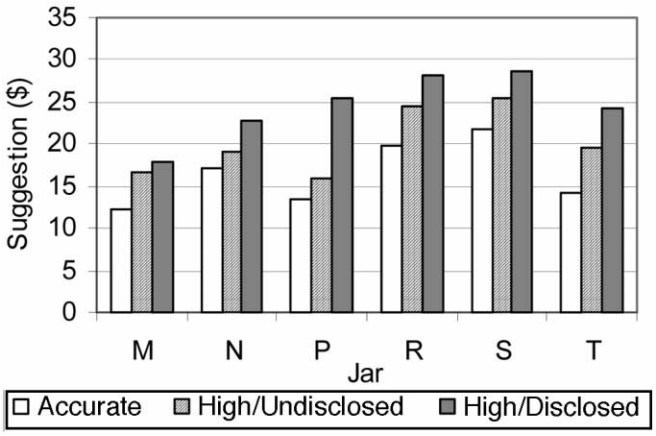

Figure 1. Advice provided for each jar, by condition

pares advisor's suggestions to actual jar values, to the mean value of advisor personal estimates of jar values in the accurate condition (we call this "virtual error"),[7] and finally to each advisor's personal estimate of jar values. The last three rows of Table 5 thereby provide three different measures of advisors' propensity to exaggerate jar values in the three conditions. Planned pairwise comparisons on each of the three measures demonstrate that advisors gave advice that was higher in the high (that is, conflict-of-interest) conditions than in the accurate condition ($p < .05$ for all three measures, by least-significant-difference test).[8]

More interesting, and as predicted, all three measures also reveal that disclosure led to greater distortion of advice. The amount that advisors exaggerated, calculated by subtracting advisors' own personal estimates from their public suggestions, was significantly greater in the high/disclosed condition than in either of the other two conditions ($p < .05$) and

7. This "virtual error" is intended to reflect how participants would have performed had it not been for the general tendency of our participants to underestimate the value of the jars. In calculating virtual error, we use advisors in the accurate condition to serve as proxies for determining an impartial subjective value of the jars, since these participants held the jars, saw the jars closer and longer than did estimators, knew a range of true values, and had no incentive to bias their valuations.

8. Because of missing data points, degrees of freedom fluctuate slightly in Tables 5, 6, and 7. This also explains, for example, why (in Table 5) the standard deviations of the means fluctuate when comparing the second through fourth rows within conditions, and (along with rounding-off errors) it explains why "advisor suggestion" minus "advisor personal estimate" is not identical to "advisor suggestion minus advisor personal estimate."

Table 4. Jar Values and Advisors' Personal Estimates in the Accurate Condition, by Jar

	M	N	P	R	S	T
Actual value ($)	10.01	19.83	15.58	27.06	24.00	12.50
Advisor personal estimate	11.85	16.73	12.75	18.39	21.30	13.07
Significance of difference (p)	<.05	<.01	<.001	<.001	<.10	.39

significantly greater by the other two measures as well: advisor suggestion minus actual jar values and advisor suggestion minus the average of personal estimates in the accurate condition ($p < .05$ for both). In the accurate condition, for example, advisors provided estimators with suggestions of jar values that were, on average, within $1 of their own personal estimates. In the high/undisclosed condition, however, advisors gave suggestions that were $3.32 greater than their own personal estimates, and in the high/disclosed condition, they gave suggestions that were inflated more than twice as much, at more than $7 above their own personal estimates. Disclosure, it appears, did lead advisors to provide estimators with more biased advice.

The first row of Table 5 shows that advisors in both the high/undisclosed and high/disclosed conditions believed that the coin jars were more valuable than did advisors in the accurate condition, although this difference was only marginally significant ($F(1, 57) = 3.43$, $p = .069$). This hints at the possibility that advisors may, to some degree, have been persuaded by their own suggestions. Perhaps convincing themselves that the jars were worth more somewhat assuaged their guilt about providing elevated estimates to estimators. Personal estimates were higher still, but not significantly so, under the high/disclosed condition than under the high/undisclosed condition.

3.2. Estimators' Estimates

Table 6 summarizes results for estimators' estimates. Estimates of jar values differed across the three experimental treatments ($F(2, 66) = 7.99$, $p < .01$). Planned comparisons revealed that estimates were higher in the two conflict-of-interest conditions than in the accurate condition ($p < .001$) and were also higher in the high/disclosed condition than in high/undisclosed condition, although this difference is not significant ($p = .19$).

As the standard deviations (S.D.s) in the first row suggest, estimator estimates were also more widely dispersed; that is, the variance of estimates was greater in the two conflict-of-interest conditions than in the

Table 5. Advisor Exaggeration of Jar Values

	Accurate (N = 24)	High/Undisclosed (N = 22)	High/Disclosed (N = 21)	Significance of Advisor Incentives (p) (Accurate versus High Conditions)	Significance of Disclosure (p) (Conflict-of-Interest Conditions)
Advisor personal estimate	15.62 (2.39)	16.79 (3.22)	16.95 (3.78)	.69	.71
Advisor suggestion	16.48 (3.50)	20.16 (4.81)	24.16 (8.40)	<.01	<.05
Advisor suggestion minus actual	−1.68 (3.50)	2.00 (4.81)	5.78 (8.51)	<.01	<.05
Advisor suggestion minus average of personal estimates, accurate condition	.75 (3.50)	4.43 (4.81)	8.35 (8.47)	<.01	<.05
Advisor suggestion minus advisor personal estimate	.82 (2.29)	3.32 (4.10)	7.10 (6.35)	<.05	<.05

Note. Standard deviations are in parentheses.

Table 6. Estimator Estimates of Jar Values

	Accurate ($N = 27$)	High/Undisclosed ($N = 26$)	High/Disclosed ($N = 27$)	Significance of Advisor Incentives (p) (Accurate versus High Conditions)	Significance of Disclosure (p) (Conflict-of-Interest Conditions)
Estimator estimate	14.21 (2.20)	16.81 (3.56)	18.14 (5.00)	<.001	.19
Estimator absolute error	5.25 (1.58)	5.14 (1.31)	6.69 (2.44)	<.363	<.01
Estimator absolute virtual error[a]	3.41 (1.36)	4.52 (1.58)	6.20 (2.62)	<.001	<.01
Absolute difference between estimate and advisor's suggestion	3.61 (3.27)	5.17 (3.33)	7.64 (5.55)	<.05	.11

Note. Standard deviations are in parentheses.

[a] Virtual error uses advisors in the accurate condition to serve as a proxy for determining an impartial subjective value of the jars by examining what would have occurred had jars been worth what accurate advisors thought, on average, the jars were worth.

accurate condition ($p < .01$) by Levene's test for the equality of variances. In addition, consistent with predictions that stem from both economic and psychological considerations, variance appears higher in the high/disclosed condition (S.D. = 5.00) than the high/undisclosed condition (S.D. = 3.56), although this difference is not statistically significant ($p = .39$).

Most important, however, and consistent with our pessimistic predictions regarding the potentially adverse effects of disclosure on recipients of advice, mean absolute estimator error is significantly greater with disclosure than without disclosure, whether measured on the basis of actual jar values ($p < .01$) or "virtual" values derived from comparisons with the mean personal estimate of advisors in the accurate condition ($p < .01$). Because of the combination of greater bias in advice and greater dispersion of estimates, estimators were less accurate with disclosure than without it.

3.3. Estimator Discounting of Suggestions

In the last row of Table 6, we see that the two conflict-of-interest conditions showed increased discounting of advisors' suggestions. In other words, the absolute difference between the suggestions given and the estimates that estimators made was greater in the two conflict-of-interest conditions than in the accurate condition ($p < .05$). However, the difference between the high/disclosed and high/undisclosed conditions was not significant ($p = .11$).

Although disclosures did increase discounting by estimators, albeit not significantly, this discounting was not sufficient to offset the increase in the bias of the advice they received. As Table 6 (fourth row) shows, estimator discounting increased, on average, less than $2 from the accurate condition to the high/undisclosed condition and less than $2.50 from the high/undisclosed condition to the high/disclosed condition. However, Table 5 (second row) shows that suggestions increased, on average, almost $4 from the accurate condition to the high/undisclosed condition and increased $4 again from the high/undisclosed condition to the high/disclosed condition. Thus, while estimators in the high/disclosed condition discounted suggestions about $4 more than did estimators in the accurate condition, the advice given in the high/disclosed condition was almost $8 higher than advice given in the accurate condition. Instead of correcting for bias, estimates were approximately 28 percent higher in the high/disclosed condition than in the accurate condition (first row of Table 6).

3.4. The Bottom Line

Table 7 summarizes payoffs in the three experimental conditions. Although significance levels vary, the basic pattern of results revealed in the table is consistent: estimators earned less money when conflicts of interest were disclosed than when they were not, and advisors made more money with disclosure than without disclosure. In addition, estimators made the most money in the accurate condition, in which there was no conflict of interest. Comparing advisors' payoffs across conditions is somewhat more problematic, owing to differences in their payoff schedules between experimental conditions.

3.5. Effects of Feedback

Recall that subjects received feedback, in the form of actual jar values, in the last three rounds. As a result, feedback about jar values was confounded with simple experience with the estimation task. Although the effects of feedback per se are difficult to assess for this reason, the results provided no grounds for concluding that either experience with the task or feedback lessened the biasing effects of disclosure. We examined the effect of feedback on overestimation of jar values using a 3 (condition) × 2 (feedback) × 3 (round) ANOVA with repeated measures on the last two factors. Neither the main effect of feedback nor its interaction with the other variables emerged as statistically significant. Not only did feedback fail to help estimators, but the trend is actually in the opposite direction, with estimators earning (nonsignificantly) less in feedback rounds 4–6 (mean = 1.35) than in nonfeedback rounds 1–3 (mean = 1.64). Advisors, in contrast, earned (nonsignificantly) more in feedback rounds 4–6 (mean = 1.75) than in nonfeedback rounds 1–3 (mean = 1.61). Given the small number of rounds, however, the conclusion that feedback does not help estimators or helps advisors should be treated with caution. Perhaps with more feedback over a much larger number of rounds, estimators would have eventually realized the extent to which their judgments were being affected by the advice given to them, and perhaps such learning would have been facilitated by disclosure.

4. DISCUSSION

A superficial analysis of disclosure helps to explain the popularity of this purported remedy for conflicts of interest: all parties appear to benefit.

Table 7. Estimator and Advisor Payoffs per Round

	Accurate	High/Undisclosed	High/Disclosed	Significance of Advisor Incentives (p) (Accurate versus High Conditions Taken Together)	Significance of Disclosure (p) (Conflict-of-Interest Conditions)
Estimator payoff	1.64 (.64)	1.59 (.65)	1.25 (.81)	.346	.052
Advisor payoff	1.69 (.66)	1.50 (.85)	1.86 (1.18)	N.A.[a]	.07
Estimator virtual payoff[b]	2.42 (.84)	1.89 (.71)	1.42 (.65)	<.001	<.05
Advisor virtual payoff[b]	2.35 (.70)	2.08 (1.17)	2.60 (1.34)	N.A.[a]	<.05

Note. Standard deviations are in parentheses.

[a] Payoffs (actual or virtual) across incentive condition are not comparable for advisors because advisors' payoff schedules are different across incentive conditions.

[b] Virtual payoffs use advisors in the accurate condition to serve as a proxy for determining an impartial subjective value of the jars by examining what would have occurred had jars been worth what accurate advisors thought, on average, the jars were worth.

However, a more complex analysis calls this optimistic appraisal into question. Disclosure, at least in the context of the admittedly stylized experiment discussed in this paper, benefited the providers of information but not its recipients. To the extent that a similar effect occurs outside the experimental laboratory, disclosure would supplement existing benefits already skewed toward information providers. In particular, disclosure can reduce legal liability and can often forestall more substantial institutional change. We do not believe that this is a general result—that is, that disclosure always benefits providers and hurts recipients of advice—but it should challenge the belief that disclosure is a reliable and effective remedy for the problems caused by conflicts of interest.

4.1. When Might Disclosure Help?

One factor that might be important in determining whether disclosure hurts or helps is the expertise of the recipients of advice. Holding all else equal, we would expect disclosure to be more effective when recipients of advice have extensive professional experience that could potentially provide them with a reasonably accurate idea about how the conflict of interest will affect their advisors. For example, government regulators may benefit from disclosure of conflicts of interest when advised by researchers who provide inputs to policy decisions. Likewise, judges (although perhaps not jurors, who have less expertise) may be better able to evaluate the claims of expert witnesses if their conflicts of interest are disclosed. The lack of impact of feedback in our experiment, however, suggests that experience is unlikely to ensure that the recipients of information will respond optimally to a disclosed conflict of interest. Moreover, even if recipients of information gain sophistication with experience, advisors are likely to also gain from experience, enabling them to provide increasingly persuasive but biased advice.

The general conclusion that disclosure is most likely to help the sophisticated estimator is somewhat dismaying, since unsophisticated estimators are exactly the ones who are most likely to need protection from exploitation. Such naïve recipients of advice would include individual investors who rely on information from stock analysts (or from auditors), individual home buyers who rely on advice from realtors, or the typical hospital patients who rely on medical professionals for advice. The paradigmatic example of the person who disclosure is unlikely to help is the medical patient who deals with only a small number of doctors, does so infrequently, lacks expertise in medicine, and enters the patient-doctor relationship trusting the doctor. This person is unlikely

to know how the physician's conflicts of interest—or the disclosure of these conflicts—is likely to influence the physician's advice, and is likely to be uncertain regarding what to do about it. A final consideration regarding the limited protection that experience can provide is the fact that many of the most significant decisions people make in their lifetimes are made only once or but a small number of times.

4.2. Potential Limitations

Like all laboratory experiments, the one reported in this paper can be criticized in terms of external validity. For example, given the small size of the stakes in our study relative to those operating in the real world, it is intuitively plausible that larger incentives would reduce the magnitude of the perverse effects we observed. Larger incentives, it might be imagined, would motivate estimators to be more cautious and to make fewer errors. Our participants, however, did appear highly motivated by the amounts of money offered to them; in fact, it was the promise of such money that induced them to participate. Participants also seemed strongly motivated by the desire to do well at the task. Furthermore, past research has failed to show that increasing financial incentives can eliminate either reliance on cognitive heuristics or the biases they produce (Camerer and Hogarth 1999; Thaler 1991). Anchoring heuristics, which may have contributed to overreliance on advice in our study, have proved particularly robust (Northcraft and Neale 1987), even when participants are explicitly motivated and instructed to avoid such heuristics (Wilson et al. 1996).

Not only estimators, but also advisors in our experiment also faced incentives that were smaller than (and somewhat different from) those faced in real-world situations. Most important, in many real-world situations, there are incentives for truth telling and honest advising, including the preservation of reputation, the solicitation of future business, and the avoidance of both criminal and civil charges in court. In our experiment, in contrast, there were no opportunities for reputation building because estimators and advisors were randomly rematched to pairs in each round, and there were no financial penalties for distorting the truth. While it is likely that advisors in our study felt some inclination to deliver useful advice when they were in the role of the advisor (see, for example, Camerer [2003] for a discussion of altruism and concerns for fairness in one-shot economic games; see also Grice [1975]), these internal motives probably underrepresented the strength of incentives for honesty in the real world. However, incentives not only for honesty

but also to inflate advice are likely to be much greater in the real world. As recent business scandals suggest, the financial benefits of manipulating consumers and investors can be sufficiently great and the chances of being prosecuted for fraud sufficiently small that, on balance, many real-world advisors also find it in their interest to provide biased advice.

Another difference between our study and at least some real-world settings is that our study examines advisory relationships in which estimators were advised by a single person for each item in question. Given that, in some settings, one purpose of disclosure may be to warn consumers to consider getting a second opinion, it would be interesting to investigate whether disclosure in fact serves such a function. There are reasons to doubt that this will be the case. For example, although patients are commonly informed of their right to obtain a second opinion, second opinions are often not sought, perhaps because patients do not want to second-guess or potentially insult their primary physicians (Foreman 2001). Moreover, when second opinions are obtained, the first opinion often carries the day. As already discussed, prior research suggests that even when an initial suggestion is totally discredited, it often continues to have an impact on those who initially hear it. Thus, the opportunity to obtain a second opinion (or even knowing exactly how much to discount advice, for example, "totally") may be insufficient to fully protect consumers from the biased advice that conflicts of interest can initially produce.

In sum, we have shown that disclosure cannot be assumed to protect recipients of advice from the dangers posed by conflicts of interest. Disclosure can fail because it (1) gives advisors strategic reason and moral license to further exaggerate their advice and (2) it may not lead to sufficient discounting to counteract this effect. The evidence presented here casts doubt on the effectiveness of disclosure as a solution to the problems created by conflicts of interest. When possible, the more lasting solution to these problems is to eliminate the conflicts of interest. As Surowiecki (2002, p. 38) commented in an article in the *New Yorker* that dealt specifically with conflicts of interest in finance, "[T]ransparency is well and good, but accuracy and objectivity are even better. Wall Street doesn't have to keep confessing its sins. It just has to stop committing them."

REFERENCES

Action on Smoking and Health. 1997. Tobacco Deal Has Potential for Even Greater Dirty Tricks. http://www.no-smoking.org/sept97/9-11-97-1.html.

Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments on Strategic Interaction*. Princeton, N.J.: Princeton University Press.

Camerer, Colin F., and Robin M. Hogarth. 1999. The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk and Uncertainty* 19:7–42.

Camerer, Colin F., George Loewenstein, and Martin Weber. 1989. The Curse of Knowledge in Economic Meetings: An Experimental Analysis. *Journal of Political Economy* 97:1232–54.

Coffee, John C., Jr. 2004. Gatekeeper Failure and Reform: The Challenge of Fashioning Relevant Reforms. Berkeley Program in Law and Economics Working Paper No. 160. University of California School of Law, Berkeley.

Crawford, Vincent P., and Joel Sobel. 1982. Strategic Information Transmission. *Econometrica* 50:1431–51.

Dana, Jason, and George Loewenstein. 2003. A Social Science Perspective on Gifts to Physicians from Industry. *Journal of the American Medical Association* 290:252–55.

Davis, Michael, and Andrew Stark. 2001. *Conflict of Interest in the Professions*. Oxford: Oxford University Press.

Dye, Ronald A. 2001. An Evaluation of Essays on Disclosure and the Disclosure Literature in Accounting. *Journal of Accounting and Economics* 32:181–235.

Foreman, Judy. 2001. First Rule: Don't Hesitate to Get Second Opinion. http://www.myhealthsense.com/F010522_secondOpinion.html.

Galinsky, Adam D., and Thomas Mussweiler. 2001. First Offers as Anchors: The Role of Perspective-Taking and Negotiator Focus. *Journal of Personality and Social Psychology* 81:657–69.

Gibbons, Robert, Frank J. Landry, Denise L. Blouch, David L. Jones, Frederick K. Williams, Catherine R. Lucey, and Kurt Kroenke. 1998. A Comparison of Physicians' and Patients' Attitudes toward Pharmaceutical Industry Gifts. *Journal of General Internal Medicine* 13:151–54.

Grice, H. Paul. 1975. Logic and Conversation. Pp. 41–58 in *Speech Acts,* edited by Peter Cole and Jerry L. Morgan. Syntax and Semantics, vol. 3. New York: Academic Press.

Hastie, Reid, David A. Schkade, and John W. Payne. 1999. Juror Judgments in Civil Cases: Effects of Plaintiff's Requests and Plaintiff's Identity on Punitive Damage Awards. *Law and Human Behavior* 23:445–70.

Hart, Philip. 1975. February 3. *Congressional Record* (Senate), 94th Cong., 1st Sess., p. 2019.

Healy, Paul M., and Krishna G. Palepu. 2001. Information Asymmetry, Corporate

Disclosure, and the Capital Markets: A Review of the Empirical Disclosure Literature. *Journal of Accounting and Economics* 31:405–40.

Issacharoff, Samuel. Forthcoming. Legal Responses to Conflicts of Interest. Chap. 13 in *Conflicts of Interest: Challenges and Solutions from Law, Medicine and Organizational Settings,* edited by Don A. Moore, Daylian M. Cain, George Loewenstein, and Max Bazerman. Cambridge: Cambridge University Press.

Lee, Fiona, Christopher Peterson, and Larissa Z. Tiedens. 2004. Mea Culpa: Predicting Stock Prices from Organizational Attributions. *Personality and Social Psychology Bulletin* 30:1636–49.

Loftus, Elizabeth F. 1979. *Eyewitness Testimony*. Cambridge, Mass.: Harvard University Press.

Mercer, Molly. Forthcoming. The Fleeting Effects of Disclosure Forthcomingness on Management's Reporting Credibility. *Accounting Review.*

Monin, Benoit, and Dale T. Miller. 2001. Moral Credentials and the Expression of Prejudice. *Journal of Personality and Social Psychology* 81:33–43.

Moore, Don A., and George Loewenstein. 2004. Self-Interest, Automaticity, and the Psychology of Conflict of Interest. *Social Justice Research* 17:189–202.

Northcraft, Gregory B., and Margaret A. Neale. 1987. Experts, Amateurs, and Real Estate. *Organizational Behavior and Human Decision Processes* 39: 84–97.

Ross, Lee, Mark R. Lepper, and Michael Hubbard. 1975. Perseverance in Self-Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm. *Journal of Personality and Social Psychology* 32:880–92.

Skurnik, Ian, Gordon B. Moskowitz, and Marcia Johnson. 2002. Biases in Remembering True and False Information: The Illusions of Truth and Falseness. Unpublished manuscript. University of Toronto.

Stark, Andrew. 2000. *Conflict of Interest in American Public Life*. Cambridge, Mass.: Harvard University Press.

Strack, Fritz, and Thomas Mussweiler. 1997. Explaining the Enigmatic Anchoring Effect: Mechanisms of Selective Accessibility. *Journal of Personality and Social Psychology* 73:437–46.

Surowiecki, James. 2002. The Talking Cure. *New Yorker*, December 9.

Thaler, Richard H. 1991. The Psychology and Economics Conference Handbook. Pp. 189–95 in *Quasi Rational Economics*, edited by Richard H. Thaler. New York: Russell Sage Foundation.

Tversky, Amos, and Daniel Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science* 185:1124–31.

U.S. Congress. House. 2001. Bipartisan Campaign Reform Act of 2001. 107th Cong., 1st sess. H. Rep. 2356.

———. 2002. Bipartisan Corporate and Auditing Accountability, Responsibility, and Transparency Act of 2002. 107th Cong., 2d Sess. H. Rep. 3763 (S. 2673).

Verrecchia, Robert E. 2001. Essays on Disclosure. *Journal of Accounting and Economics* 32:97–180.

Wegner, Daniel M. 1994. Ironic Processes of Mental Control. *Psychological Review* 101:34–52.

Wilson, Timothy D., and Nancy Brekke. 1994. Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations. *Psychological Bulletin* 116:117–42.

Wilson, Timothy D., Christopher E. Houston, Kathryn M. Etling, and Nancy Brekke. 1996. A New Look at Anchoring Effects: Basic Anchoring and Its Antecedents. *Journal of Experimental Psychology: General* 125:387–402.