Exp Econ (2008) 11: 154–173
DOI 10.1007/s10683-007-9164-2

Learning under supervision: an experimental study

Raghuram Iyengar · Andrew Schotter

Received: 14 July 2006 / Revised: 5 February 2007 /
Accepted: 12 February 2007 / Published online: 14 July 2007

Abstract In many market environments, for example in investment banking, sales-force management and others, workers and supervisors work closely as a team. Workers are paid a fixed salary and supervisors determine any raises, which are typically dependent on how well the organization does. In such scenarios, a supervisor who constantly offers suggestions can create a problem—typically a worker cannot ignore his supervisor's advice, yet if such advice is wrong and is followed, it will only decrease firm profits. We conduct a laboratory experiment to address a question critical for such settings—does the relationship between advisor and worker interfere with the learning abilities of the worker? The answer is a resounding no. In fact, subjects who have a supervisor advising them and whose advice is costly to ignore actually learn better than those with an advisor whose advice can be ignored. An even more striking result is that advisees as well as advisors in both these conditions learn better than subjects with no advisors. Our result can be attributed to the presence of advice and has direct relevance to learning in many environments.

Keywords Group decision making · Advice · Learning · Supervision

JEL Classification C91 · C92

Electronic supplementary material The online version of this article (http://dx.doi.org/10.1007/s10683-007-9164-2) contains supplementary material, which is available to authorized users.

R. Iyengar (✉)
Wharton School, 3730 Walnut Street, 700 Jon M. Huntsman Hall, Philadelphia, 19104 PA, USA
e-mail: riyengar@wharton.upenn.edu

A. Schotter
New York University, 269 Mercer St. 7th Floor, New York, NY 10003, USA
e-mail: andrew.schotter@nyu.edu

1 Introduction

The Principal-Agent theory, while relevant for some contracts in the workplace, fails to encompass many of the relationships that can exist between a worker and his supervisor[1] (Perrow 1986; Eisenhardt 1989; Sujan et al. 1994; Kohli et al. 1998). Many times, the incentives of the worker and the supervisor are somewhat aligned and the situation is akin to a team relationship (Radner and Marschak 1972; Cooper and Kagel 2005; Kocher and Sutter 2005; Cox and Hayne 2006). A common example of such a relationship, from investment banking, sales force management and others, is one wherein a worker is paid a fixed salary and any raises are determined by his supervisor who works closely with the worker. Typically these raises are a function of how well the organization does as a whole. In such settings, however, if the supervisor offers suggestions, a problem is created for the worker—he cannot completely ignore the advice (even when he knows it to be wrong) for fear of displeasing the boss and causing him to lower the raise. The worker is caught between pleasing his boss and making what he feels are the best decisions. Such a trade-off can affect worker performance and learning-by-doing. We explore the dynamics of such tradeoffs within a laboratory experiment wherein advisors and advisees make repeated decisions.

We ask a question that is critical for such settings—does the relationship between worker (agent) and supervisor (advisor) interfere with the learning abilities of the worker? In other words, are intrusions by the advisor so distracting that the agent fails to learn the essentials of his job?[2] Surprisingly the answer is a resounding no. In fact, subjects in our experiment who have bosses advising them actually learn better than those with no laboratory bosses. This result is true even when the given advice can be ignored with no cost. Interestingly, we find that advisors also learn better than subjects with no bosses. The reason, we conjecture, is that advice forces both advisors and advisees to think deductively about the problem. This reflection of the problem is different from the more inductive trial and error process that people typically go through. Past research can give an indication of why such reflection might occur in the presence of advice.

The work on accountability suggests that when people are accountable, they engage in a more effortful, self-critical search (Lerner and Tetlock 1994; Tetlock 1983). This search can lead participants to pay greater attention to the information that they use and also develop a greater awareness of their cognitive processes (Lerner and Tetlock 1999). In our setting, we have advisors who are giving advice and advisees who are evaluating the given advice and then making a decision. The reward structure (discussed later) is such that both are affected by each other's actions. Such a setting makes both advisors and advisees accountable for their actions. It is this accountability that can lead subjects to reflect on the problem more and gain a better assessment of the situation.

[1]We use supervisor, advisor and boss interchangeably in this paper. Similarly, we use agent, advisee and worker interchangeably.

[2]Overbearing parents who smother their children with directives also have to fear that they may stunt the ability of their children to learn how to function independently and make good decisions on their own.

An increase in reflection and its benefits have been pointed out in past research on problem solving and learning. For instance, Chi et al. (1989) found that good learners reflected on the problem in a way that went beyond the information that was given. In contrast, bad learners merely considered just the given information and did not make any other inferences. Such self-explanation and abstraction has also been found to improve learning in many domains (Ferguson-Hessler and de Jong 1990; Pressley et al. 1992). Hence, we surmise that it is the process of accountability and reflection that enhances learning in the presence of advice.[3] Interestingly, past work has in fact shown that subjects, when acting alone, have a difficult time learning in environments where decisions are repeated and there are rewards at the end of every round (Merlo and Schotter, 1999, 2003). However, it is precisely this type of environment that exists when people function in markets. Hence, our results speak towards the beneficial aspect of advice for learning within market environments.

Our work adds to the previous research on decision-making within small groups (Cooper and Kagel 2005; Bornstein and Yaniv 1998; Bornstein and Gneezy 2002; Kocher and Sutter 2005; Cox and Hayne 2006). For instance, Cooper and Kagel (2005) investigate whether individuals or two-person teams do better in signaling games. They find that teams play more strategically than individuals and trace this improvement to the dialogue between teammates. Similar to their result, we find that both advisors and advisees learn better than those subjects with no laboratory bosses. In addition, we also find that the improvement in learning within teams occurs even without any overt conversation between the team members. In a similar vein, Kocher and Sutter (2005) compare individual and group behavior in beauty contest games and find that teams learn faster than individuals. Recently, Cox and Hayne (2006) also investigated whether small groups are rational in their decision making and if such decision making is systematically different from that carried out by individuals. Interestingly, they find that five-person teams that had face to face discussions and used endogenously determined rules performed worse than individual decision makers. They attribute this result to excess of information that is present within groups. In our setting, however, both team members have access to the same level of information while making decisions.

Our research also adds to past work on advice and its effects (Krishna and Morgan 2001; Ottaviani and Sorensen 2003). Both Krishna and Morgan (2001) and Ottaviani and Sorensen (2003) consider the issue of expert advice. Krishna and Morgan consider a situation where experts, who are perfectly informed, offer advice to decision makers. These decision makers can affect the welfare of others. They show that in the case of advice from just one expert, the expert withholds information. Ottaviani and Sorensen consider cheap talk behavior from experts. In our experiments, we do not have experts. Finally, our work augments past research on supervisor-worker relationships and appraisal within an organizational setting (Lounamaa and March 1987; Prendergast 1993; Duarte et al., 1993, 1994; Athey et al. 2000) and in particular, to the research in marketing that has considered the relationship between sales

[3]This result is similar to that reported by Roberto Weber (2003) where he allows subjects to play the Nagel 2/3rd's (Nagel 1995) guessing game repeatedly without any feedback information. He finds there that behavior is more convergent toward the equilibrium despite the fact that no information is received. The explanation, however, is the same: forcing people to reflect on their behavior leads them to treat the problem they face differently.

personnel and their supervisors (John and Weitz 1989; Jaworksi and Kohli 1991; Oliver and Anderson 1994; Kohli et al. 1998). Typically field settings have been used to investigate the effect of supervisor-worker relationship on several dependent variables such as job satisfaction and favoritism within an organization. As our interest is in the specific issue of learning within a supervisor-worker dyad, a tightly controlled experimental setting is more suitable.

In this paper we proceed as follows. In Sect. 2, we describe the problem motivating our experiment. Section 3 outlines the experiment. In Sect. 4, we present the results. Section 5 concludes.

2 Decision problem

In our experiments, subjects played a simple game against a computerized opponent.[4] In each round, the subject had to choose a number, e, between 0 and 100 called their decision number. They were told that their computerized partner would always choose the number 37. After this number is chosen, a random number is independently generated from a uniform distribution over the interval $[-q, +q]$ for both the subject and his computerized opponent. These numbers (the decision number and the random number) are then added together and a "total number" is defined for each the real and computerized players. Payoffs are determined by comparing the total numbers of the real and computerized subjects and awarding the real player a fixed payoff of M if his total is larger than that of the computerized opponent. If his total number is smaller, then he receives a payoff of m, $m < M$. The cost of the decision number chosen is given by a convex function $\mathrm{c(e)} = \mathrm{e}^2/r$, where r is a constant. This amount is then subtracted from these fixed payments to determine a subject's final payoff. Hence, in these experiments, there is a trade-off in the choice of decision numbers: higher numbers generate a higher probability of winning the big prize but also imply a higher decision cost. By letting $r = 500$, $q = 40$, $M = 29$ and $m = 17.2$, and holding the computerized player's choice fixed at 37, our subjects face a rather simple decision problem with a quadratic payoff function whose peak is at 37.[5]

We feel that this particular form of decision task is suitable for our purpose: it presents subjects with a complete information maximization problem for which optimal actions can be calculated, yet such a problem is still complicated enough that a deductive solution would be out of the grasp of most subjects. This being the case, subjects should be exploring the strategy space (numbers from 0 to 100) and it is exactly this process that is of interest. In addition to the reasonable complexity, the problem is also simple to describe. This ensures that subjects clearly understand the problem that they are facing.

3 Experimental procedures

Subjects were brought into a computer lab and were all handed a common set of instructions. The instructions had the description of the entire experiment, the roles

[4]These experiments are similar to those in Merlo and Schotter (1999).

[5]See Bull et al. (1987) for a version of this problem where both subjects are real players.

played by the advisor and the agent during the experiment and determination of their respective payoffs. Subjects were called decision makers of type P and type A and not Advisors or Agents. After the subjects read the instructions, they were randomly paired up in groups of two. One member of the group was assigned to the role of the advisors (we use advisors and subjects of type P interchangeably in this paper) while the other was the agent (we use agents and subjects of type A interchangeably in this paper). Thereafter, the groups were escorted to computers where they began the experiments. We ensured that the decisions made by one group were not visible to others by seating them at sufficient distance from each other.

The agents repeated the experiment for 75 rounds. Before every round, the advisors wrote down what they thought is the right number for their agents to choose. The advisor's suggestion would be his advice to the agent. There was absolutely no other form of communication between the advisor and the advisee, i.e., they were not allowed to speak, or make any verbal or facial expressions to each other. They were, in fact, not even allowed to react to the feedback they were getting from the computer. We strictly enforced complete silence during the entire experiment.

The payoffs in each round depended on the treatment run: in the Disagreement Cost treatment, the agents received a payoff from the decision problem in round t, π_t, which was composed of three parts. First, they received a fixed payment of either M or m depending upon whether their total number was above or below that of the computer. Next, from that fixed payment a decision cost $\mathrm{C}(\mathrm{e}_{it}) = \mathrm{e}_{it}^2/500$ was deducted and finally an amount equal to $0.025(\mathrm{e}_{it} - \mathrm{a}_{it})^2$ was further deducted, where e_{it} is the decision number chosen by subject i in round t and a_{it} is the advice he received from the advisor in that round. This last term is a disagreement cost specifying a cost to disagreeing with the advisor. In the no-cost treatment this last term was set to zero so there was no cost to following or ignoring an advisor's suggestion. The advisor's payoff was equal to $3/4$'th of their agent's payoff. This payoff was inclusive of their disagreement cost so the advisors also lost money if their agent did not follow their advice. Thus, the incentives of the advisor and the advisee were aligned.

After each round of the experiment both subjects received some information, which was shown on the computer screen. In the Disagreement Cost treatment, both subjects saw the decision number chosen, the advice given, the advice cost, the payoff (inclusive of decision costs) of type-A subjects before the advice cost was deducted, the net payoff (inclusive of both decision costs and advice costs) to type-A subjects and the net payoff to the type P subjects (3/4'th of type-A's net payoff). In the No-disagreement cost treatment, there was no cost of disagreeing and the net payoff of the agent included only the decision cost. Thus, in this condition, both subjects saw the decision number chosen, the advice given, the net payoff to type-A subjects (inclusive of decision cost) and the payoff to type-P subjects (3/4'th of the net payoff to type-A subjects). Thus, at all times in our experiment, both team members had access to the same information.

After the 75 rounds of the experiment were over we administered a "surprise quiz" (Merlo and Schotter 1999). To do this we first asked the advisors to leave the room. The remaining subjects, i.e., the agents, were told that they would be participating in one more round of the experiment which involved making the same decision that they had made for the past 75 rounds but with a few changes. One, there was no advice going to be given for this round (i.e., no type P subjects or advisors were present). Sec-

Table 1 Experimental design

Treatment	Description	Number of pairs	Data
1 (NA)	Agent alone	24 subjects	MS (1999)
2 (DC)	Agent + advisor disagreement cost	19 pairs	Current paper
3 (NDC)	Agent + advisor No disagreement cost	14 pairs	Current paper

ond, this round was worth 75 times the value of the previous rounds. In other words, this one choice was worth exactly as much as the sum of the 75 previous choices. Thus, since they were playing for quite big stakes in this "surprise-quiz round", the choice should serve as a sufficient statistic of how much the agents learned during the past 75 rounds. After the agents made their surprise quiz decision, they were escorted out of the room, given a questionnaire, paid and let go. The advisors were then ushered into the room and given the same surprise quiz, paid and let go. Hence, when we ask questions about how well our subjects learned over the course of the experiment we will be comparing the results of their surprise quiz choices over different treatments.

The treatment used for comparisons are the two treatments just described as well as one other reported in Merlo and Schotter (1999) (MS 1999). In MS (1999) the experiment performed was identical to that performed here except that the type-A subject (agent) acted alone without the presence of an advisor (i.e., without the advice from a type-P subject). While the experiment in MS (1999) was similar, it was run in a different time period. This could potentially create a problem while comparing its results with those reported in this paper. However, the instructions given to the subjects in MS (1999) were similar to those given to the type-A subjects in the current experiments. In addition, the subject pool for our experiments as well as those in MS (1999) is also very similar. This should alleviate some of the concerns with using the results of MS (1999) as a benchmark. Table 1 lists the three treatments just described.

3.1 The optimal action

Given the computerized partner's choice of 37 and the uniform distributions for the random terms, our decision makers face a one-person decision problem in which they compete against a player who always chooses 37. Using the joint distribution of the random terms, we can define an expected payoff function for our type-A subjects (agents) that takes the following form[6]:

$$E\pi_t = \alpha + \beta_1 \mathrm{e}_t + \beta_2 \mathrm{e}_t^2 - k(\mathrm{e}_t - \mathrm{a}_t)^2,$$

where e_t is the choice made by the type-A subjects and a_t is the advice offered by his type-P pair member at time t and k is a known constant. In this expression, the first

[6] See Schotter and Weigelt (1992) for a full derivation.

three terms constitute the payoff function associated with the decision problem while the last term is the disagreement cost function. The agent must trade-off his need to satisfy his boss's desires and the desire to maximize his payoff from the decision problem he faces. When k equals 0, we go to the No-disagreement cost case.

Given the parameters of the experiments, i.e., $r = 500$, $q = 40$, $M = 29$ and $m = 17.2$, the objectively true (but unobservable) coefficients of the expected payoff function are $\alpha = 18.94$, $\beta_1 = 0.079$ and $\beta_2 = -0.0011$. In other words, this is the expected payoff function that each of our subjects did in fact face. Note that when $k = 0$, we get the optimal decision to be, $\mathrm{e}_t = -\frac{\beta_1}{2(\beta_2)} = -\frac{0.079}{2(-0.0011)} \cong 37$.

As these coefficients are unobservable, at any time t during the experiment the best the subject can do is to use the data from the past $t-1$ rounds to estimate the values of α, β_1 and β_2 in his payoff function (k is known with certainty at the start of the experiment). He can then use the estimated coefficients, $\hat{\alpha}_t$, $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$, to decide on the optimal choice of e_t by simply maximizing the estimated payoff function, $E\hat{\pi}_t$. Taking the derivative of $E\hat{\pi}_t$ with respect to e_t and equating it to zero, we find:

$$\frac{dE\hat{\pi}_t}{d\mathrm{e}_t} = \hat{\beta}_{1t} + 2\mathrm{e}_t\hat{\beta}_{2t} - 2k\mathrm{e}_t + 2k\mathrm{a}_t = 0,$$

$$\mathrm{e}_t^{\text{optimal}} = -\frac{\hat{\beta}_{1t}}{2(\hat{\beta}_{2t} - k)} - \frac{2k\mathrm{a}_t}{2(\hat{\beta}_{2t} - k)}.$$

The second order conditions are satisfied if $\hat{\beta}_2 < 0$ or $\hat{\beta}_2 < k$. If these conditions fail then we get optima at the corners, which is either 0 or 100. Note that when $k = 0$ the optimal e_t is simply the choice that, given the current estimates for the coefficients of the payoff function, maximizes the subject's payoff without any disagreement-cost. With $k > 0$, the optimal e_t is one that optimally trades off an agent's need to maximize his payoff against his disagreement cost given the advice.

4 Results

We present our results in two parts. In part one we analyze the surprise quiz round choices of our subjects in the three treatments. This is the data we use to make statements about the ability of our subjects to learn in these three environments. In part two, we discuss the advice-giving and advice-following behavior of our subjects.

4.1 Surprise quiz behavior

Tables 2a, 2b and 2c present the choices of our subjects in the surprise quiz rounds in each of the three treatments. Recall that in the MS (1999) experiment subjects performed our task alone without an advisor. The mean and median surprise quiz choices of subjects in this condition were 51 and 50 respectively (see Table 2c).

A median test rejects the hypothesis that 37 is the median of the MS (1999) agent choices. The message of MS (1999) was that people fail to learn appropriately when they are in a setting in which the task is repeated often and there is a reward at the end of each round. The question is whether the presence of advice can alleviate this problem.

Table 2a Surprise quiz choices. Disagreement cost treatment

Pair	Surprise quiz choice—Type A	Surprise quiz choice—Type P
1	40	47
2	38	50
3	38	38
4	37	38
5	23	38
6	64	51
7	23	29
8	48	47
9	65	38
10	52	51
11	23	20
12	60	60
13	25	17
14	23	82
15	19	3
16	1	1
17	50	25
18	21	1
19	36	1
Mean	36.10	33.52
Median	37	38

In our experiments, we find a remarkable result: the process of giving and receiving advice enhances the learning ability of both advisees (type-A) and advisors (type-P). For example, in the Disagreement Cost condition, the mean and median choices were 36.1 and 37 for the type-A subjects and 33.5 and 38 for the type-P subjects, respectively (Table 2a). Neither of these medians are significantly different from 37—median tests show that ($p \leq 1$) for the type-A agents and ($p \leq 0.648$) for the type-P subjects. These results are striking as, with the presence of a disagreement cost for any deviations, it is not obvious that both advisees and advisors would learn better as compared to subjects in a condition wherein there are no advisors. For the No-Disagreement Cost experiment, the situation is slightly different. Here the mean and median choices were 31.28 and 39.5 for the type-A subjects and 43.35 and 43 for the type-P subjects (Table 2b). Only the type-A subjects had a median that was not significantly different from 37. For the type-P subjects we had to reject that hypothesis at the 2% level. This indicates that advisors, whose advice could be ignored, did not learn as well as those advisors whose advice could not be ignored. Yet, these type-P advisors still learned better than subjects in the MS (1999) experiment who actually did the experiments—a Wilcoxon test indicates that we can reject the hypothesis that the sample of type-P surprise quiz rounds came from the same population as those of the subjects in MS (1999) ($p \approx 0$).

Table 2b Surprise quiz choices. No-disagreement cost treatment

Pair	Surprise quiz choice—Type A	Surprise quiz choice—Type P
1	37	50
2	39	45
3	40	44
4	10	38
5	10	63
6	1	37
7	1	20
8	50	40
9	66	57
10	1	1
11	40	50
12	61	80
13	42	42
14	40	40
Mean	31.2	43.35
Median	39.5	43

We also find that the overall distribution of choices differed when advice was and was not available. If we compare the choices of subjects in MS (1999) with those from subjects in the experiments with advice, we find that the former distribution has a much larger variance and a greater selection of high decision numbers. A set of Pearson's Chi-Squared tests support this conclusion by reporting significant differences between all experiments with and without advice.[7]

We find another important result—advice diminishes the number of subjects who choose dominated strategies. For example, in the surprise quiz rounds of all three experiments, any choice of 65 or more is dominated by choosing 0 (Schotter and Weigelt 1992). While in MS (1999) 10 out of 24 subjects chose a dominated strategy in their surprise-quiz round, in our No-Disagreement Cost experiment only 1 type-P subject and 1 type-A subject made dominated choices. In other words there were only 2 out of 28 such choices. Even in the Disagreement Cost experiment, only 1 type-P subject made surprise-quiz choices strictly greater than 65. This is a very strong difference and indicates that these subjects clearly learned a minimum lesson.

The key finding then is that learning is fostered when advice is given even if there is no cost to ignore it.

4.2 Advice giving and following

The surprise quiz data presented a snapshot of what was learned by our subjects. It said nothing, however, of the process they went through. In this section of the paper

[7] NA and DC(type-A) ($p \approx 0$), NA and DC(type-P) ($p \leq 0.0002$), NA and NDC(type-A) ($p \leq 0.04$), NA and NDC(type-P) ($p \leq 0.07$).

Table 2c Surprise quiz: MS (1999)

Subject	Surprise quiz choice
1	65
2	45
3	100
4	77
5	0
6	45
7	41
8	68
9	70
10	65
11	45
12	35
13	0
14	44
15	68
16	0
17	70
18	69
19	50
20	50
21	45
22	50
23	30
24	100
Mean	51.33
Median	50

we analyze the advice giving and advice following strategies of our subjects. We ask several questions. First, was advice followed by our type-A subjects? Second, was the advice offered to them accurate in the following sense—if we assume that our type-P subject was an amateur (or professional) econometrician and knew that the payoff function was quadratic so that it was fully defined by three coefficients, α, β_1 and β_2. At each round t, upon observing the choices made up to round $t-1$, they could use the choice-payoff data to get an estimate of each of these three coefficients, $\hat{\alpha}_t$, $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$, and calculate the payoff maximizing choice. They could then offer that payoff maximizing choice as their advice. At any time t, this should yield an advice of $\mathrm{a}_t = -\frac{\hat{\beta}_{1t}}{2\hat{\beta}_{2t}}$ as this is the period-t best guess about the peak of the payoff function. We ask whether such a model explains the advice giving behavior of our type-P subjects.

Finally, as the optimal choice for a type-A subject is to trade-off his decision payoff against his disagreement cost, we can ask if this is what type-A subjects do. In other words, similar to the type-P subjects, at any time during the experiment our type-A subjects should have an estimate of the payoff-optimal decision, $e_t = -\frac{\hat{\beta}_{1t}}{2\hat{\beta}_{2t}}$. Given the advice, a_t, they should modify the decision and choose $e_t = -\frac{\hat{\beta}_{1t}}{2(\hat{\beta}_{2t}-k)} - \frac{2ka_t}{2(\hat{\beta}_{2t}-k)}$. We ask whether this process furnishes a good model of type-A advice-following behavior.

4.3 Advice following

To answer whether type-A subjects follow advice, we regressed their 75 decision choices on the advice they were offered. This model is termed as the advice model. We estimated the parameters of this model using OLS but, as the data is of a time series nature, there is a possibility of autocorrelation of the errors. This autocorrelation can cause a bias in the OLS estimated standard deviations. To address this issue, we used the Newey–West (Newey and West 1987) estimator. The Newey–West estimator provides a robust, consistent estimator for the covariance in the presence of autocorrelated disturbances with an unspecified covariance structure. The advice model is specified as:

$$e_{it} = \gamma + \zeta a_{it} + \varepsilon_t,$$

where, e_{it} is the decision made by subject i in period t, a_{it} is the period t advice received by the subject, and ε_t is a disturbance term, which can be correlated with the previous disturbances. Tables 3a and 3b show the regression results for the Disagreement Cost as well as the No-Disagreement Cost condition, respectively.

The results for the Disagreement Cost condition show that the advice model is extremely well specified. For example, 75% of the subjects have R^2's of 0.69 or greater for these simple regressions while the median R^2 is 0.82. In addition, about 25% of the subjects have an R^2 of 0.98 or more. Clearly this implies that this simple model does a good job of explaining behavior.

The results for the No-Disagreement Cost condition are worse. Here the mean R^2 is 0.59 while 75% of the subjects have R^2's greater than only 0.34 (only 25% have R^2's of 0.80 or more).[8] Obviously, subjects ignored the advice more when it was free to do so. However, as we have seen, as a group they still learned far better than those subjects (in MS 1999) who did not receive advice at all. This is an interesting result and it reinforces the idea that it is the process of advice giving and receiving, and not the type of advice per se, that makes people think differently about the problem.

An alternative to the advice model posited above is that subjects optimally tradeoff payoff maximization versus disagreement cost minimization by choosing $e_t = -\frac{\hat{\beta}_{1t}}{2(\hat{\beta}_{2t}-k)} - \frac{2ka_t}{2(\hat{\beta}_{2t}-k)}$. We have to be careful here as the first order conditions for a maximum are only necessary conditions and, given any $t-1$ period history, it is possible that $\hat{\alpha}_t$, $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$ defines a local minimum instead. In such cases, the best

[8] We also pooled the data across the two conditions and regressed the decision on advice. We find that there is no significant difference in the intercepts across the two conditions. There is, however, a significant difference in the coefficient of advice: the no-disagreement cost group has a smaller coefficient as compared to that in the disagreement cost group. This is consistent with the results shown in Tables 3a and 3b, where we find that subjects in the disagreement cost condition follow advice more closely as compared to those in the no-disagreement cost condition. The estimation results using pooled data is available from the authors upon request. We thank one of the reviewers for suggesting this analysis.

Table 3a Advice model—regression results. Advice following: disagreement cost treatment

Pair	$\hat{\gamma}$	$\hat{\xi}$	R^2
1	−0.02	1.04	0.98
	(1.05)	(0.017)	
2	5.36	0.93	0.82
	(3.23)	(0.07)	
3	4.52	0.82	0.70
	(1.88)	(0.08)	
4	0	1.00	1.00
5	7.99	0.82	0.34
	(16.7)	(0.382)	
6	4.54	0.89	0.79
	(2.22)	(0.05)	
7	2.19	0.96	0.93
	(1.60)	(0.03)	
8	25.97	0.44	0.57
	(5.25)	(0.12)	
9	20.02	0.66	0.80
	(3.08)	(0.05)	
10	3.49	0.91	0.11
	(10.49)	(0.25)	
11	5.37	0.98	0.70
	(3.07)	(0.08)	
12	12.77	0.82	0.69
	(6.47)	(0.09)	
13	2.19	0.96	0.93
	(1.60)	(0.03)	
14	11.04	0.6	0.56
	(3.57)	(0.08)	
15	0	1	1
16	0.03	0.97	0.89
	(0.02)	(0.02)	
17	8.96	0.81	0.89
	(1.93)	(0.05)	
18	0.71	0.97	0.99
	(0.23)	(0.008)	
19	1.96	0.97	0.98
	(0.99)	(0.02)	

advice is at the corners of the feasible set, either 0 or 100. As time evolves, however, these estimates can change rapidly with the estimated payoff function changing repeatedly from being concave to convex and back again.

Table 3b Advice model—regression results. Advice following: no-disagreement cost treatment

Pair	$\hat{\gamma}$	$\hat{\zeta}$	R^2
1	20.49	0.43	0.58
	(2.45)	(0.06)	
2	12.57	0.66	0.61
	(3.39)	(0.09)	
3	−0.13	0.96	0.21
	(8.62)	(0.22)	
4	6.74	0.64	0.16
	(5.55)	(0.17)	
5	28.17	0.45	0.20
	(7.07)	(0.13)	
6	2.00	0.90	0.37
	(9.23)	(0.59)	
7	0.95	0.99	0.95
	(1.01)	(0.02)	
8	29.25	0.24	0.34
	(2.35)	(0.06)	
9	22.25	0.69	0.59
	(5.95)	(0.09)	
10	−0.83	0.80	0.61
	(1.29)	(0.07)	
11	12.31	0.59	0.43
	(4.17)	(0.05)	
12	0.00	1.00	1.00
13	−0.33	1.01	0.94
	(1.16)	(0.03)	
14	−0.42	0.98	0.83
	(1.49)	(0.03)	

To alleviate this problem we consider the behavior of only those subjects (we call them "concave A-type subjects") who, during the last 20 rounds, consistently had estimated payoff functions that were concave and hence had defined interior maxima. We look at the behavior of these subjects only and restrict the analysis to the last 20 rounds of the experiment. An analysis of the last 20 rounds leaves sufficient number of rounds for any learning to occur and for the shape of the payoff function to not fluctuate wildly from period to period.

By regressing the choices of our concave type-A subjects on their optimal decision given their disagreement costs, we specify the following model (optimal-action model):

$$\mathrm{e}_{it} = \gamma + \zeta \hat{\mathrm{e}}_{it}^{*} + \varepsilon_t,$$

where $\hat{\mathrm{e}}_{it}^{*}$ is the period-t estimated best choice for agent A, and ε_t is a white noise disturbance term. In other words, $\hat{\mathrm{e}}_{it}^{*}$ is that choice that is best for the concave Type-A subject given his estimates of the coefficients of the payoff function and his disagreement costs.[9] Tables 4a and 4b show the regression results for the Disagreement Cost as well as the No-Disagreement Cost condition respectively.

These results show that the advice model outperforms the optimal action model. We get further corroboration when we estimate the optimal action model after pooling data from subjects in disagreement cost condition for the last 20 rounds and find that the overall R^2 for this model is 0.45 whereas the R^2 for the advice model is 0.77. We find a similar result (i.e., the advice model outperforms the optimal action model) within the no-disagreement cost condition as well.

In summary, it appears as if subjects follow advice rather closely in both of our experiments although they do so more often when it is costly not to deviate.

4.4 Advice giving

Now that we know that advice is basically followed, an obvious next question to ask is how good is this advice. One model suggests that advice is informationally optimal i.e., at any time t, the type-P subjects offer the advice that is their best guess of the peak of the payoff function based on their current information set. For the same reasons as stated above, we only look at concave type-P subjects (these coincide with their concave type-A counterparts as we use the same criterion).

We regress the advice offered by concave type-P subjects over the last 20 rounds of the experiment on the estimated peak of their payoff function and test to see if its coefficient is significantly different from 1.0. Tables 5a and 5b show the estimated coefficients and the results of the hypothesis test.

From the tables it is evident that most concave type-P advisors are offering advice that is informationally optimal. In the costly advice condition, we cannot reject the null hypothesis of the right-hand side coefficient equal to 1 for 11 out of the 12 concave advisors. In the no-cost advice treatment, there were only 2 out of the 7 concave type-P advisors for whom we could reject the null hypothesis.

5 Conclusions

In several business environments, workers and supervisors have common interests and collaborate as a team. Workers are paid a fixed salary and supervisors determine any raises, which are typically dependent on how well the organization does. In such scenarios, a supervisor who constantly offers suggestions can create a problem—typically a worker cannot ignore his supervisor's advice, yet if such advice is wrong and is followed, it will only decrease firm profits. In this paper, we investigated agent learning in such supervisor-worker dyads within a laboratory experiment.

[9] For the optimal action model, we wanted the optimal decision (e^*_{it}) to be stable and not change rapidly from one time period to another. Thus, we settled on using the data from only the last 20 rounds of the concave Type-A subjects. We also analyzed the data for the other subjects (non-concave subjects) and found that their optimal decision was either 0 or 100 and these optimal decisions fluctuated a lot. Thus, the data from these subjects will not reveal much into how they are trading off the payoff with the disagreement cost.

Table 4a Optimal action model—regression results. Advice following: disagreement cost treatment

Pair	Category	$\hat{\gamma}$	$\hat{\xi}$	R^2
1	good	44.42	0.18	0.02
		(14.70)	(0.28)	
2	bad			
3	good	25	0	1
4	bad			
5	bad			
6	good	61.81	−0.29	0.06
		(12.70)	(0.27)	
7	good	15.39	0.49	0.21
		(7.12)	(0.22)	
8	good	13.63	0.65	0.23
		(13.40)	(0.27)	
9	good	52.99	0.11	0.01
		(16.80)	(0.27)	
10	good	−3.93	1.11	0.86
		(4.49)	(0.10)	
11	good	1.59	0.93	0.55
		(3.80)	(0.20)	
12	good	35.43	0.46	0.15
		(16.30)	(0.25)	
13	bad			
14	good	49.31	−0.03	0.01
		(13.50)	(0.25)	
15	bad			
16	bad			
17	good	9.30	0.68	0.16
		(13.51)	(0.36)	
18	bad			
19	good	−0.45	1.07	0.63
		(8.91)	(0.19)	

Good refers to Concave Type-A subjects

Bad refers to subjects that are not Concave Type-A

We considered two treatments—a disagreement cost condition wherein workers were penalized for moving away from the given advice and a no-disagreement cost condition where they were free to ignore the advice. We found that agents in both the

Table 4b Optimal action model—regression results. Advice following: no disagreement cost treatment

Pair	Category	$\hat{\gamma}$	$\hat{\xi}$	R^2
1	bad			
2	bad			
3	bad			
4	good	−1.09	1.20	0.07
		(37.50)	(1.01)	
5	bad			
6	good	22.81	0.37	0.01
		(24.85)	(0.87)	
7	bad			
8	bad			
9	good	183.79	−2.91	0.41
		(33.72)	(0.81)	
10	bad			
11	good	−4.54	0.98	0.01
		(48.27)	(1.85)	
12	good	−76.23	3.83	0.18
		(67.55)	(1.92)	
13	good	60.27	−0.39	0.05
		(19.05)	(0.39)	
14	good	57.72	−0.39	0.08
		(16.30)	(0.31)	

Good refers to Concave Type-A subjects

Bad refers to subjects that are not Concave Type-A

No-Disagreement Cost treatment and the Disagreement Cost condition learned better than subjects in a treatment wherein advice was absent. Interestingly, even the advisors in these advice treatments learned better than agents in no-advice treatment. The results suggest that the presence of advice forces both advisors and advisees to think deductively about the problem. This reflection of the problem leads to insights that may not be gleaned by a subject involved in a learning task in which he is constantly getting payoff feedback. Such a stimulus-response atmosphere tends to lead subjects to act myopically as opposed to sitting back and seeing the forest for the trees. However, it is precisely these type of environments that exist when people function in markets and make choices repeatedly which are reinforced by an immediate payoff. Thus, our results have direct relevance to the beneficial aspect of advice for learning within market environments.

There are several ways of extending our framework to address other issues in group decision making. We have considered a situation where a supervisor's con-

Table 5a Advice giving: disagreement cost treatment

Pair	Category	$\hat{\gamma}$	$\hat{\zeta}$	R^2	Null hypothesis $\zeta = 1$
1	good	19.37	0.79	0.06	Accept Null
		(31.96)	(0.73)		
2	bad				
3	good	25	0	1	Reject Null
4	bad				
5	bad				
6	good	34.13	0.34	0.00	Accept Null
		(95.40)	(2.32)		
7	good	26.41	0.15	0.00	Accept Null
		(34.46)	(1.49)		
8	good	−8.71	0.90	0.02	Accept Null
		(82.40)	(1.38)		
9	good	−105.52	3.12	0.06	Accept Null
		(154.80)	(2.92)		
10	good	−24.93	1.62	0.18	Accept Null
		(34.77)	(0.82)		
11	good	−20.44	1.49	0.11	Accept Null
		(25.66)	(0.99)		
12	good	−75.06	2.47	0.09	Accept Null
		(99.61)	(1.76)		
13	bad				
14	good	13.69	0.93	0.09	Accept Null
		(106.87)	(2.91)		
15	bad				
16	bad				
17	good	−45.08	1.89	0.21	Accept Null
		(36.07)	(0.85)		
18	bad				
19	good	−216.45	6.77	0.10	Accept Null
		(186.03)	(4.79)		

Good refers to Concave Type-P subjects

Bad refers to subjects that are not Concave Type-P

Table 5b Advice giving: no disagreement cost treatment

Pair	Category	$\hat{\gamma}$	$\hat{\zeta}$	R^2	Null hypothesis $\zeta = 1$
1	bad				
2	bad				
3	bad				
4	good	1.59	0.90	0.27	Accept Null
		(13.10)	(0.35)		
5	bad				
6	good	13.28	0.74	0.11	Accept Null
		(14.04)	(0.49)		
7	bad				
8	bad				
9	good	−102.85	3.79	0.08	Accept Null
		(120.44)	(2.90)		
10	bad				
11	good	−55.51	3.49	0.06	Accept Null
		(82.49)	(3.16)		
12	good	−76.23	3.84	0.18	Accept Null
		(67.55)	(1.92)		
13	good	60.27	−0.39	0.05	Reject Null
		(19.05)	(0.39)		
14	good	42.25	−0.06	0.14	Reject Null
		(2.03)	(0.03)		

Good refers to Concave Type-P subjects

Bad refers to subjects that are not Concave Type-P

stant advice creates a problem for the worker, yet they are working as a group with common interests. In addition, both supervisor and worker have the same information for making a decision. There could be other situations where the role asymmetry (i.e., one member is a worker and another is a supervisor) might lead to information asymmetry and this can have an effect on decisions. We can extend our framework to address this issue. One experimental set up could be to have an intergenerational game where the workers of one game, after gaining some experience, become the supervisors in the following game. Another interesting issue that can be explored is the effect of the frequency of advice on decisions. In our current set up, the supervisor offers advice at the beginning of every round. One extension can be to investigate the effect of advice if it is offered after every 5 or 10 rounds. A second extension would be to let the members choose when to offer and receive advice. We plan to address some of these issues in future work to develop more insights into the effect of advice in group decision making.

Acknowledgements Support for this research was offered by the National Science Foundation Grant Number SES 0111640 as well as the Center for Experimental Social Science (C.E.S.S.) and the C.V. Starr Center for Applied Economics at New York University. The authors would also like to thank Maher Said and Mahati Guttikonda for their research assistance and Tom Palfrey and the members of the (C.E.S.S.) Experimental Lunchtime Seminar for their comments. The first author also thanks Gal Zauberman for insightful discussions.

References

Athey, S., Avery, C., & Zemsky, P. (2000). Mentoring and diversity. *The American Economic Review*, *90*, 765–786.

Bornstein, G., & Gneezy, U. (2002). Price competition between teams. *Experimental Economics*, *5*, 29–38.

Bornstein, G., & Yaniv, I. (1998). Individual and group behavior in the ultimatum game: are groups more "rational" players?. *Experimental Economics*, *1*, 101–108.

Bull, C., Schotter, A., & Weigelt, K. (1987). Tournament and piece rates: an experimental study. *Journal of Political Economy*, *95*, 1–33.

Chi, M., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: how students study and use examples in learning to solve problems. *Cognitive Science*, *15*, 145–182.

Cooper, D. J., & Kagel, J. H. (2005). Are two heads better than one? Team versus individual play in signaling games. *American Economic Review*, *95*(3), 477–509.

Cox, J. C., & Hayne, S. C. (2006). Barking up the right tree: are small groups rational agents?. *Experimental Economics*, *9*(3), 209–222.

Duarte, N. T., Goodson, J. R., & Klich, N. R. (1993). How do i like thee? Let me appraise the ways. *Journal of Organizational Behavior*, *14*, 239–249.

Duarte, N. T., Goodson, J. R., & Klich, N. R. (1994). Effects of dyadic quality and duration on performance appraisal. *The Academy of Management Journal*, *37*, 499–521.

Eisenhardt, K. (1989). Agency theory: assessment and review. *Academy of Management Review*, *14*, 57–74.

Ferguson-Hessler, M., & de Jong, T. (1990). Studying physics texts: differences in study processes between good and poor solvers. *Cognition and Instruction*, *7*, 41–54.

John, G., & Weitz, B. (1989). Salesforce compensation: an empirical investigation of factors related to use of salary versus incentive compensation. *Journal of Marketing Research*, *16*, 1–14.

Jaworksi, B. J., & Kohli, A. K. (1991). Supervisory feedback: alternative types and their impact on salesperson's performance and satisfaction. *Journal of Marketing Research*, *28*, 190–201.

Kocher, M., & Sutter, M. (2005). The decision maker matters. Individual versus team behavior in experimental beauty-contest games. *Economic Journal*, *115*, 200–223.

Kohli, A. K., Shervani, T. A., & Challagalla, G. N. (1998). Learning and performance orientation of salespeople: the role of supervisors. *Journal of Marketing Research*, *35*, 263–274.

Krishna, V., & Morgan, J. (2001). A model of expertise. *Quarterly Journal of Economics*, *116*, 747–775.

Lerner, J. S., & Tetlock, P. E. (1994). Accountability and social cognition. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (Vol. 1, pp. 3098–3121).

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*(2), 255–275.

Lounamaa, P. H., & March, J. G. (1987). Adaptive coordination of a learning team. *Management Science*, *33*, 107–123.

Merlo, A., & Schotter, A. (1999). A surprise-quiz view of learning in economic experiments. *Games and Economic Behavior*, *28*, 25–54.

Merlo, A., & Schotter, A. (2003). Learning by not doing: an experimental investigation of observational learning. *Games and Economic Behavior*, *42*, 116–136.

Nagel, R. (1995). Experimental results on interactive competitive guessing. *American Economic Review*, *85*, 1313–1326.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*, 703–708.

Oliver, R. L., & Anderson, E. (1994). An empirical test of the consequences of behavior and outcome-based sales control systems. *Journal of Marketing*, *58*, 53–67.

Ottaviani, M., & Sorensen, P. N. (2003). Professional advice: the theory of reputational cheap talk. Working paper.

Perrow, C. (1986). *Complex organizations*. New York: Random House.

Prendergast, C. (1993). A Theory of 'yes men'. American Economic Review, 757–770.

Pressley, M., Wood, E., Woloshyn, V., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: attempting to construct explanatory answers facilitates learning. *Educational Psychologist*, *27*, 91–109.

Radner, R., & Marschak, J. (1972). *Economic theory of teams*. New Haven: Cowles Foundation and Yale University Press.

Schotter, A., & Weigelt, K. (1992). Asymmetric tournaments, equal opportunity laws and affirmative action: some experimental results. *Quarterly Journal of Economics*, *108*, 511–539.

Sujan, H., Weitz, B. A., & Kumar, N. (1994). Learning orientation, working smart and effective selling. *Journal of Marketing*, *58*, 39–52.

Tetlock, P. E. (1983). Accountability and complexity of thought. *Journal of Personality and Social Psychology*, *45*, 74–83.

Weber, R. (2003). Learning with no feedback in a competitive guessing game. *Games and Economic Behavior*, *44*, 134–144.