Underpredicting Learning after Initial Experience with a Product

DARRON BILLETER
AJAY KALRA
GEORGE LOEWENSTEIN

For products that require skills to use, such as computers, cell phones, and sports equipment, consumers' purchase and usage decisions often depend on their prediction of the speed with which they will master the relevant skills. In this article, we identify a systematic pessimism in predictions of such skill learning occurring in the initial skill-acquisition phase of product use. After initially trying new products, people underpredict how quickly they will acquire the skills required for product use. Further, we find that this underprediction of learning is due to a failure to appreciate how rapidly task experience leads to a shift from system 2 to system 1 processing. In six experiments, we document the effect, examine its generality across several tasks, and demonstrate its consequences for product devaluation and abandonment. We conclude with a discussion of implications for customer service, promotions, and the design of new products.

Since the seminal investigation of the diffusion of the hybrid corn seed innovation in Iowa (Ryan and Gross 1943), the adoption of new products has been examined from a variety of perspectives. Factors governing the adoption of new products have been examined in economics (Katz and Shapiro 1986; Tirole 1988), strategy (Leonard-Barton 1992), information systems (Venkatesh and Davis 2000), health care (Budman, Portnoy, and Villapiano 2003), and marketing (Rogers 1976). The primary focus of the marketing research has been on firm strategies and product characteristics that facilitate product adoption. Among the major findings is that two of the more important factors influencing the adoption of a new product are ease of use and perceived usefulness (Davis 1989; Davis, Bagozzi, and Warshaw 1989).

Bagozzi, Davis, and Warshaw (1992) find that intentions to try a new technology depend on consumer attitudes toward the process of learning and expected reactions to success and failure. Similarly, Thompson, Hamilton, and Rust (2005) show that consumers focus on product capabilities before trying a product but on ease of learning after they begin using the product. One category of goods for which initial product experience is particularly important is skill-based products that can only be fully used and appreciated after consumers acquire the requisite skills (Burson 2007; Murray and Haubl 2007). Examples include sporting goods (e.g., skis and sailboards), do-it-yourself products (e.g., home improvement and furniture), and many electronic devices and appliances (e.g., computers, cameras, and bread makers). Not only the initial adoption but also the continued use of such products is likely to depend on consumers' expectations regarding their own ability to master usage of the product.

There is some evidence that initial learning often serves

Darron Billeter is an assistant professor at the Marriott School of Management, Brigham Young University, Provo, UT 84602 (darron@byu.edu). Ajay Kalra is a professor of marketing at the Jones Graduate School of Business, Rice University, Houston, TX 77005 (ajay.kalra@rice.edu). George Loewenstein is the Herbert A. Simon Professor of Economics and Psychology, Social, and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA, 15213 (gl20@andrew.cmu.edu). The authors would like to thank Marcel Goic, Sharad Borle, Joachim Vosgerau, Linda Argote, and Zachary Burns for their help. The authors also thank the editor, area editor, and four reviewers for their helpful suggestions. This material is supported by the National Science Foundation grant EEEC-0540865, a research grant from the Center for Analytical Research at Carnegie Mellon University, a grant from the Pennsylvania Department of Health's Commonwealth Universal Research Enhancement Program, and a PhD grant funded by Richard Staelin at Carnegie Mellon. This work is based on the first author's doctoral dissertation.

John Deighton served as editor and Steve Hoch served as associate editor for this article.

Electronically published August 2, 2010

 0093-5301/2011/3705-0002$10.00. DOI: 10.1086/655862

as a barrier to new product adoption. A recent doctoral dissertation reports that consumers spend an average of only 20 minutes trying to operate new electronics items before they give up and that 50% of products returned to electronics stores that consumers claim to be defective are actually fully functional (den Ouden et al. 2006). There is also considerable anecdotal evidence that consumers either discard new products or do not fully use them. One survey that examined UK consumers' usage of newly purchased kitchen items found that between 60% and 72% of consumers purchasing yogurt makers, plastic-bag-sealing devices, juicers, and coffee machines ultimately failed to make use of their acquisitions (esure Insurance and ICM 2006). Another survey of 500 people found that 22% of respondents did not learn how to use a high-technology gift they had received in the past year (Walker 2007).

In this article, we investigate how consumers form perceptions of their own ability to acquire skills. Our focus is on tasks novel to the consumer that require skill development and for which performance improves with repetition. We contrast consumers' expectations of their own speed of skill mastery with their actual speed of skill mastery and find that consumers make a systematic error after initial product use. Specifically, before hands-on experience with a task, consumers are overconfident about both their initial mastery of a task and their speed of learning. However, this overconfidence quickly transmutes to underconfidence when consumers begin the skill-acquisition process. We find that this tendency to underpredict learning persists during the early stages of learning but that calibration eventually improves with experience.

We demonstrate the generality of this effect across tasks that span visuospatial and fine motor skills, using different measures such as unit performance and time to complete the task. We investigate behavioral consequences of the effect, finding that product valuations are dynamically affected, with initial product valuations declining after product trial when consumers lose confidence in their own ability to master usage of the product.

Self-Assessments of Future Performance

Although accurate assessments of performance allow consumers to make better purchasing decisions (Alba and Hutchinson 2000), most research finds that consumers are not very good at predicting their own future performance. Thus, for example, one meta-analysis of the relationship between self-perceptions of knowledge and actual performance found correlations ranging from .47 for athletics (motor skills) to .17 for interpersonal skills (Mabe and West 1982; Morwitz 1997).

A well-established finding is that people overpredict their performance in many domains (Dunning, Heath, and Suls 2004; Epley and Dunning 2006; McGraw, Mellers, and Ritov 2004), including using product features (Meyer, Zhao, and Han 2008). People tend to overrate themselves when it comes to both predicting their own absolute performance (e.g., the planning fallacy; Buehler, Griffin, and Ross 1994) and predicting relative performance (e.g., the better-than-average effect; Kruger and Mueller 2002). Evidence of underconfidence is rarer, although several studies have found that people tend to be underconfident on difficult tasks (Kruger 1999; Moore and Cain 2006; Moore and Kim 2003), which is known as the hard/easy effect (Lichtenstein, Fischhoff, and Phillips 1982).

Another exception to the general finding of overconfidence in predicting absolute performance—and the article that is most closely related to our work—is research by Koriat, Sheffer, and Ma'ayan (2002) focusing on memory tasks. In a series of studies, they exposed subjects to a series of word pairs, had them predict their own likelihood of recalling the test word when shown the cue word, and then measured their actual recall. Some of these word pairs were well known and already familiar to the participant (e.g., cow and milk); others were not (e.g., fox and citizen). On the first trial of this task, the authors observed reliable overconfidence; subjects thought they would be more likely to remember the test words than they subsequently were able to. In later trials, however, subjects underestimated their own subsequent recall of test words, a regularity the authors dubbed the "underconfidence-with-practice" effect. Koriat et al. (2006) proposed that participants are initially overconfident because they put too much weight on the familiar word pairs in making their predictions. After experiencing the task, however, Koriat et al. posit, participants become focused disproportionately on the unfamiliar word pairs that tripped them up during the first trial and, as a result, exhibit underconfidence.

Skill Acquisition

A consistent finding in the literature on skill acquisition is that learning is initially rapid, decelerates with experience, and loosely conforms to a power function (Newell and Rosenbloom 1981). This power law of learning has been demonstrated for a wide range of tasks (Newell and Rosenbloom 1981) and skill-acquisition categories, such as auditory, visual, and perceptual motor skills from Fleishman's (1975) classic taxonomy of human performance.

Improvements in performance are generally accompanied by, and undoubtedly partly driven by, qualitative changes in cognitive processes. Most of the research addressing cognitive processes underlying task learning is in agreement that learning involves a transition from more deliberate, conscious, effortful information processing to more automatic, unconscious, and effortless processing. For example, Fitts (1964) proposed a still widely accepted three-stage framework for understanding skill acquisition in which the learner first gathers the facts needed to understand and perform the task, next begins to try out the task, and finally begins to master the task, at which point the learner's actions are fast, smooth, accurate, effortless, and largely removed from the learner's awareness (Evans 2007). Anderson (1982) likewise proposes a three-stage model of skill acquisition that is remarkably similar (a point noted by Ackerman [1987] in his review of the literature on task learning). Despite differences in specific details, both theoretical frameworks posit a transition from processing that is relatively slow, difficult, serial, inefficient, effortful, deliberative, conscious, and controlled to processing that is relatively quick, effortless, efficient, automatic, nonconscious, and characterized by an ability to engage in multitasking behavior. These are precisely the characteristics of the system 1 and system 2 modes of information processing introduced by Stanovich and West (2002; see also Kahneman and Frederick 2005).

Not surprisingly, given the dramatic changes in cognitive processes that have been identified with the process of skill acquisition, parallel changes have been observed in the neural mechanisms associated with task engagement in different stages of the learning process (Evans 2007). For example, Haier et al. (1992) quantified the cerebral glucose metabolic rate (GMR) in Tetris video game players and found that after 4–8 weeks of daily practice, GMR in cortical surface regions decreased, despite a sevenfold increase in performance. Haier's findings support the notion that, with experience, there is a reorganization of active brain areas. Later experiments with more sophisticated brain-imaging techniques have shown that during initial novel skill learning, the control network (a network of discrete regions of the brain that control goal processing, attention, and decision making) is initially very active but becomes less active with experience (Chein and Schneider 2005; Hill and Schneider 2006; Schneider and Chein 2003).

Predicting Future Performance

Predicting one's own future performance at a task can be viewed as a form of intrapersonal perspective taking. As just described, engagement with a task tends to produce changes in mental structures that result in improvements in performance. Predicting future performance, therefore, entails taking the perspective of oneself at a time in the future when one's mental structures will be different.

There is ample evidence that people are not very good at this type of task and, in fact, make systematic errors. Research on the hindsight bias (Fischhoff 1975) shows that after individuals learn that something has happened, they have difficulty simulating their own perspective in the absence of such learning. Research on the "curse of knowledge" (Camerer, Loewenstein, and Weber 1989) shows that people have difficulty taking the mental perspective of other, less knowledgeable, people. Research on the hot-cold empathy gaps shows that people have difficulty simulating their own or other people's feelings and behavior in a different emotional state than the one they are currently in (Loewenstein 1996). All of this research suggests that people tend to project their own thoughts and feelings on the past and future as well as on other people, resulting in what has been termed "projection bias" (Loewenstein, O'Donoghue, and Rabin 2003).

Applied to learning, projection bias implies that someone who is unskilled at a task will have difficulty imagining being skilled. When system 2 processing is active during initial skill learning, we posit that it is very difficult for a person to imagine performing the task in a state in which system 1 processing dominates. This should lead learners to predict that future execution of the task will require more effort than it actually will. Conversely, once expertise has developed and system 1 processing occurs, consumers are likely to find it difficult to recall the level of effort that is required for skill acquisition. For example, Hinds (1999) finds that experts make greater errors than people with intermediate skill levels when predicting novice performance, positing that the error results from the experts' failure to recall the initial difficult period of mastery.

Consistent with previous research (Dunning et al. 2004; Epley and Dunning 2006), we predict that in the information-gathering stage (before experience), people will exhibit overconfidence. We predict, however, that after experiencing the slow, effortful processing arising from initial experience with a task, people will lower their predictions—even to the extent of underpredicting their subsequent speed of learning. At a practical level, moreover, we predict that this decline in optimism about task mastery will lead to quitting behavior and reduced valuations for products that require skill acquisition.

EXPERIMENT 1

In experiment 1 we investigate individuals' perceptions of their ability to acquire a skill and test the hypothesis that, after initial exposure to a task, people lower their performance predictions, switching abruptly from overconfidence to underconfidence.

Method and Procedure

Forty-eight participants from a paid subject pool consisting of both students and nonstudents were recruited for a show-up fee of $4 and performance-based payments. Subjects learned the classic mirror-tracing task (Snoddy 1926), which was selected because it requires acquisition of a new skill but is simple enough that the learning rate is rapid. The mirror-tracing task requires subjects to draw a shape using only a mirror's reflection. Subjects were asked to trace an unbroken line between the boundaries that were formed by two five-pointed stars, with one star placed in the interior of the other star.

Participants were first given 2 minutes to view a folder containing instructions on performing the mirror-tracing task. The instructions included the stimuli and evaluation procedures. A correctly completed trace required that the participants not cross the inner or outer boundary of the star pattern while drawing. If the boundary was touched, it was considered an error, and participants were instructed to discard the trace and begin a new drawing immediately. Subjects were instructed to correctly trace as many stars as possible in four rounds of 5 minutes each.

Immediately after viewing the instructions, participants were asked to predict the number of correct traces they would be able to complete in the four rounds. These will be called before-experience (P_{BE}) predictions. After completing the before-experience predictions, participants were given 2 minutes to try the task. At the end of this initial experience period, participants again made performance predictions for the four rounds. These will be termed after-experience (P_{AE}) predictions.

After these steps, the rounds commenced. All participants in a session began concurrently, and a buzzer signaled the completion of each 5-minute round. At the completion of each round, the researcher and the subject counted and recorded the number of correctly traced stars. This process was repeated for each of the four rounds. After each round, participants made predictions for the remaining rounds. After the four rounds ended, participants were asked to roll a die to receive payment (as described below). Finally, subjects were paid, debriefed, and dismissed.

The payment incentives were designed to ensure that subjects exerted maximum effort and made accurate predictions. Subjects were instructed that they would be paid for each round on the basis of either performance or prediction accuracy. At the end of the task, a die was rolled for each of the four rounds. If the number rolled was between one and five, payment was based on performance, which was 25 cents per trace completed. If a six was rolled, payment was based on both prediction accuracy and performance. Specifically, the formula used was \$.25 × (number of traces completed − |number of traces predicted − number of traces completed|). The incentive scheme was designed to be easily explained and incentive compatible and to ensure that subjects would be motivated to forecast their own performance as accurately as possible while exerting full effort on the task. Subjects earned an average of \$16.97 in the experiment.

We present the results using the following notations: A_n represents actual performance, where $n = 1, \ldots, 4$ indicates the round, and before-experience and after-experience predictions are indicated by P_{BEn} and P_{AEn}, respectively, where n indicates the round predicted. Predictions for later rounds are indicated by P_{jn}, where j represents the time period when the prediction is made and n is the round predicted.

Results and Discussion

There is considerable discussion about the appropriate functional form to model learning (Heathcote, Brown, and Mewhort 2000), with the power and exponential functions receiving the most attention. For all of the studies reported in this article in which it was possible to do so, we estimated the parameters from the power function ($Y = ax^b$) and the exponential function ($Y = ae^{bx}$) and compared the model fit by investigating the explained variance (r^2) captured by each model. The results in table 1 show that the power function ($r^2 = .99$) provides a better fit to the aggregate data of experiment 1 than does the exponential function ($r^2 = .97$), and the table data show a similar, consistent pattern.

TABLE 1

COMPARISON OF POWER VERSUS EXPONENTIAL FIT

	Power function ($Y = ax^b$)			Exponential function ($Y = ae^{bx}$)		
Experiment	a	b	r^2	a	b	r^2
1	6.00	.77	.99	5.21	.31	.97
2, control condition	142.50	−1.41	.99	317.41	−.82	.96
2, debias condition	151.30	−1.43	.99	340.81	−.83	.97
3	8.41	.76	.99	7.25	.31	.98
6	5.90	.71	.99	5.11	.29	.98

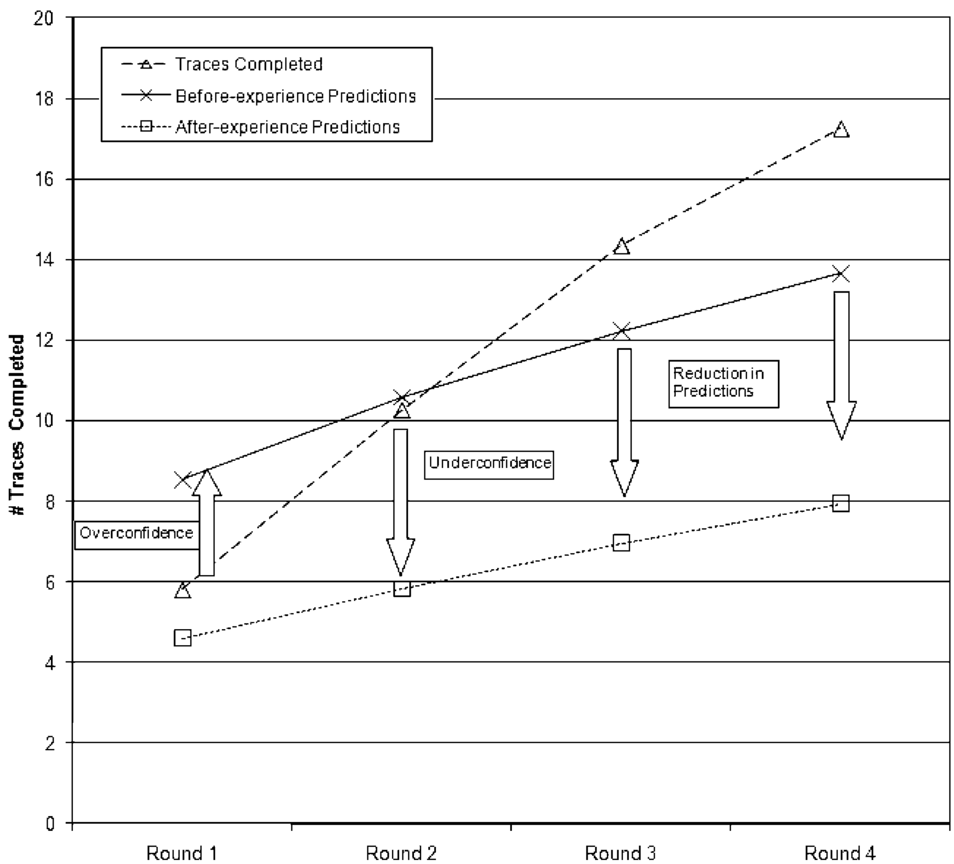
Reduction in Predictions after Experience. We compared the predictions made before and after the participants experienced the task. As can be seen in figure 1 and table 2, after acquiring experience, participants lowered their predictions for all subsequent rounds. We conducted a repeated-measures ANOVA with time (before experience, after experience) and mode (performance vs. prediction) as within-subjects variables. Comparing before- and after-experience predictions indicates a significant reduction in outlook after experience ($F(1, 47) = 23.72$, $p < .01$). Thus, moving from the information-gathering stage to the trial stage of learning, there was an immediate and broad reduction in participant outlook.

Current Predictions. We examined predictions of performance in a round made immediately preceding the predicted round. Before trying the task, participants exhibited overconfidence and significantly overpredicted their performance ($M[P_{BE1} - A_1] = 2.73$; SD = 7.94; $F(1, 47) = 5.67$, $p < .05$). After experience, the overconfidence turned to underconfidence, and participants significantly underpredicted performance ($M[P_{AE1} - A_1] = -1.21$; SD = 3.65; $F(1, 47) = 54.02$, $p < .001$). Subjects continued to underpredict their performance before rounds 2 and 3. Only at round 4 did calibration improve to the point at which the prediction error was minimal and not statistically significant.

Rate-of-Learning Predictions. The decision to persevere with a new task or product is likely to be contingent on predictions of the rapidity with which a desired level of accomplishment will be achieved. We therefore examined slope predictions made for the rate of change between the upcoming round and the final round (e.g., $M[P_{BE4} - P_{BE1}]$), comparing it to the actual rate of learning (e.g., $M[A_4 - A_1]$). Before their initial experience, participants already significantly underpredicted their rate of learning ($M[A_4 - A_1] - M[P_{BE4} - P_{BE1}] = 6.33$; $F(1, 47) = 37.78$, $p = .001$), and after experience the magnitude of inaccuracy increased ($M[A_4 - A_1] - M[P_{AE4} - P_{AE1}] = 8.10$; $F(1, 47) = 151.52$, $p < .001$). Participants continued to underpredict their rate of learning before rounds 2 and 3. The results indicate a systematic underconfidence in predicting learning. Predictions before experience already underestimate actual learning but underestimate even more severely after initial experience. Additionally, the results for this and all experiments are replicated using Bayesian analysis. The analysis is available upon request of the authors.

FIGURE 1

MIRROR-TRACING EXPERIMENT



EXPERIMENT 2

The main objective of experiment 2 is to address the robustness, with respect to the method of eliciting expectations, of the effects observed in experiment 1. Participants predict task completion times, rather than their performance in a time period unit. One issue with the previous experiment is that the scale used to elicit responses is likely to produce overconfidence if initial performance is poor (e.g., if performance is zero knots, then overconfidence is the only type of inaccuracy that can be obtained initially). An advantage of using time as a dependent measure is that since initial performance is relatively slow, the first dependent measure is a large number (as it takes many seconds to complete the task). This addresses an important type of response error known as a scale-end effect (Juslin, Winman, and Olsson 2000), which has been used to explain overconfidence and which may have been a contributor to the effects obtained in experiment 1.

Method and Procedure

Eighty-two undergraduate students participated in the experiment for a combination of extra course credit and performance-based payment. The task involved folding T-shirts. Participants were taught the task using a 40-second instructional video that demonstrates a novel way to fold a T-shirt that employs four steps for successful folding and is much quicker than the method that most people normally employ. The task required acquiring both insight and motor skills.

The experimental procedures were similar to those in the mirror-tracing experiment, with a few differences. The practice period was divided into two 40-second phases. Participants watched the instructional video twice, first without being allowed to touch the T-shirt, then with permission to begin practicing folding the T-shirt. The task required was to fold two shirts in each of five rounds. Participants timed themselves.

Each session consisted of between two and six participants seated facing their own computer terminals. To their immediate left was a flat empty work space where participants folded the T-shirts. Headsets were provided so that participants would have the flexibility to simultaneously fold the T-shirts and listen to the instructional video.

The compensation scheme was again designed to motivate

TABLE 2

RESULTS OF EXPERIMENT 1 (MIRROR-TRACING TASK)

	Round 1	Round 2	Round 3	Round 4	Slope
Traces completed	5.81 (5.95)	10.27 (8.41)	14.35 (10.24)	17.25 (10.74)	11.44 (6.51)
Prediction:					
Before experience	8.54 (7.53)	10.58 (9.10)	12.23 (10.00)	13.65 (11.04)	5.11 (4.93)
After experience	4.60 (4.41)	5.83 (5.15)	6.96 (6.09)	7.94 (6.83)	3.34 (3.49)
Before round 2		8.06 (7.37)	9.25 (8.45)	10.04 (8.81)	1.98 (2.25)
Before round 3			13.21 (10.37)	14.77 (11.72)	1.56 (2.45)
Before round 4				17.33 (11.89)	

NOTE.—Standard deviations are in parentheses.

subjects to perform maximally but also to be incentive-compatible for time predictions. As in the mirror-tracing task, participants rolled a die at the end of the experiment for each round and were paid on the basis of their performance or prediction accuracy. If the die rolled was between one and five, they were paid 1,000/(number of seconds it took to fold two shirts) in cents. If the die rolled was a six, they were paid on the basis of their prediction accuracy. Participants were paid 1,000/(number of seconds it took to fold two shirts − |predicted number of seconds it took to fold two shirts − number of seconds it took to fold two shirts|). On average, participants were paid $1.38.

The experiment is a between-subjects, single-factor design with two conditions: control and debias. In the debias condition, we informed participants of the hypothesized prediction errors. After watching the instructional video, participants in the debias condition were told, “When we conducted this study in the past, we have consistently found two things. First, we found that before people practice folding the T-shirt, they predict that they will do much better than they actually do. Second, once they start folding the T-shirts, they predict that they will do worse than they actually do.” In the debias condition, as a test that they had understood the information (i.e., as a kind of manipulation check), participants were asked to circle the correct word to complete the following two statements: (1) “before they start practicing, people predict that they will do better/worse than they actually do” and (2) “after they start folding the T-shirts, people predict they will do better/worse than they actually do.” Two subjects failed to complete both statements correctly and were dropped from the analysis.

Results and Discussion

The results for the control and debias conditions are reported in tables 3 and 4, respectively. As in experiment 1, we found that the power function provided a better fit, in each condition, to the data than did the exponential model (see table 1).

Reduction in Predictions after Experience. Consistent with experiment 1, predictions were significantly lower after experience ($F\ (1,38)_{\text{control}} = 7.52$, $p < .01$; $F\ (1,40)_{\text{debias}} = 5.96$, $p < .05$).

Current Predictions. Consistent with experiment 1, participants were overconfident before trying the task ($M\ [P_{\text{BE1}} - A_1]_{\text{control}} = 91.19$; SD $= 199.70$; $F(1, 38) = 8.13$, $p < .01$; $M\ [P_{\text{BE1}} - A_1]_{\text{debias}} = 94.22$; SD $= 166.12$; $F\ (1,40) = 13.19$, $p < .001$). Unlike experiment 1, after initial experience (attempting to fold the T-shirt while watching the instructional video the second time), participants remained overconfident ($M[P_{\text{AE1}} - A_1]_{\text{control}} = 82.96$; SD $=$ 195.80; $F\ (1,38) = 7.00$, $p < .01$; $M\ [P_{\text{AE1}} - A_1]_{\text{debias}} =$ 85.07; SD $= 20.35$; $F(1, 40) = 11.17$, $p < .01$). However, after gaining further experience in round 1, they became underconfident in predicting round 2 ($M\ [P_{22} - A_2]_{\text{control}} = -11.12$; SD $= 29.12$; $F\ (1,38) = 5.69$, $p < .05$; $M\ [P_{22} - A_2]_{\text{debias}} = -19.74$; SD $= 64.33$; $F\ (1,40) = 3.86$, $p <$ .05), and the same pattern occurred after round 2 when predicting round 3 ($M[P_{33} - A_3]_{\text{control}} = -5.27$; SD $= 9.81$; $F\ (1,38) = 11.25$, $p < .01$; $M\ [P_{33} - A_3]_{\text{debias}} = -8.69$; SD $= 30.34$; $F\ (1,40) = 3.37$, $p < .05$). We speculate that the overconfidence persisted because, unlike in experiment 1, participants did not have enough time to repeatedly attempt the task in the practice period. Participants became accurate in predicting task learning before rounds 4 and 5.

Rate-of-Learning Predictions. Before their initial experience, participants significantly underpredicted their rate of learning ($M\ [A_5 - A_1] - M\ [P_{\text{BE5}} - P_{\text{BE1}}]_{\text{control}} =$ 95.90; $F(1, 40) = 9.32$, $p < .01$; $M[A_5 - A_1] - M[P_{\text{BE5}} - P_{\text{BE1}}]_{\text{debias}} = 101.49$; $F(1, 40) = 15.31$, $p < .001$). After experience, the magnitude of the learners’ pessimism about their rate of learning increased slightly, although not significantly, with participants continuing to significantly underpredict their own speed of learning.

Comparison between Control and Debias Conditions. There were no significant differences between the control and the debias conditions in predictions of current or next period learning or in the reduction of predictions after initial experience. There was no evidence that the debiasing intervention led to any improvement in participants’ prediction accuracy.

EXPERIMENT 3

In this experiment, we returned to the mirror-tracing task and introduced a stronger debiasing manipulation.

TABLE 3

RESULTS OF EXPERIMENT 2 (T-SHIRT FOLDING TASK) CONTROL CONDITION

	Round 1	Round 2	Round 3	Round 4	Round 5	Slope
Actual time	144.60 (195.68)	43.16 (35.60)	30.52 (26.87)	27.11 (15.36)	22.70 (12.16)	121.90 (193.02)
Prediction:						
Before experience	53.41 (30.59)	44.67 (25.78)	36.69 (21.81)	31.62 (18.99)	27.41 (16.39)	26.00 (19.22)
After experience	61.64 (42.43)	53.00 (36.66)	44.03 (28.91)	40.03 (26.52)	36.72 (24.96)	24.92 (24.16)
Before round 2		54.28 (36.69)	43.85 (26.63)	37.79 (21.88)	34.41 (19.98)	19.87 (23.62)
Before round 3			35.79 (28.29)	28.33 (13.18)	25.25 (11.87)	9.54 (19.34)
Before round 4				26.62 (14.71)	24.05 (13.90)	2.56 (4.93)
Before round 5					23.64 (12.90)	

NOTE.—Standard deviations are in parentheses.

Method and Procedure

Twenty-five undergraduate students from a midwestern university participated in the experiment for extra course credit and performance-based pay. The experiment is a single-factor, within-subjects design. The mirror-tracing task from experiment 1 was selected as the learning task.

A similar procedure to experiment 1 was followed, with a few exceptions. First, after making their initial (before-experience) predictions, participants were told, "We found that before people practice mirror tracing, they predict that they will do much better than they actually do. Once they start mirror tracing, they predict that they will do worse than they actually do." Then participants were asked again to predict the number of traces for each round. After initially experiencing the task and completing their predictions, participants were reminded that "people predict that they will do worse than they actually do." Again participants were asked to make predictions. Thus, before the initial experience, participants make two sets of predictions (before and after reading about the effect), and after experiencing the task, participants make two further sets of predictions (before and after being reminded of the effect). To differentiate between the two before-experience predictions, we refer to the preinformation predictions as the control predictions ($P_{\mathrm{BE}nc}$) and the after-information predictions as the debias predictions ($P_{\mathrm{BE}nd}$). A similar incentive structure to experiment 1 was employed, except that participants were paid \$.10 instead of \$.25 for a correctly completed trace (with an average payment of \$6.70 per participant). Similar to previous experiments, we found that the power function provided a better fit to the data than did the exponential function (see table 1).

Results—Control Predictions

Reduction in Predictions after Experience. Participants significantly reduced their predictions after initial experience (see table 5). The main effect comparing before-experience and after-experience predictions was significant ($F\ (1, 24)\ =\ 15.24$, $p\ <\ .01$), indicating that after-experience predictions were more pessimistic.

Current Predictions. Before experience, participants were initially overconfident in their mirror-tracing ability; they predicted that they would, on average, complete 12.60 traces in the first round but actually completed only 8.28 ($M[P_{\mathrm{BE1c}} - A_1] = 4.32$; SD = 10.76; $F(1, 24) = 4.03$, $p < .05$). Investigation of current prediction contrasts for each round indicated that after experience ($M[P_{\mathrm{AE1c}} - A_1] = -2.96$; SD = 5.44; $F(1, 24) = 7.42$, $p < .01$) and before round 2 ($M\ [P_{22} - A_2] = -4.48$; SD = 3.37; $F(1, 24) = 44.23$, $p < .001$), participants significantly underpredicted their performance. Before rounds 3 and 4, participants only marginally underpredicted their performance. Not surprisingly, given the additional experience gained, participants became more accurate in their next period performance with greater experience.

TABLE 4

RESULTS OF EXPERIMENT 2 (T-SHIRT FOLDING TASK) DEBIAS CONDITION

	Round 1	Round 2	Round 3	Round 4	Round 5	Slope
Actual time	152.90 (172.40)	47.90 (49.90)	32.40 (27.46)	24.87 (13.93)	22.09 (11.90)	130.81 (168.86)
Prediction:						
Before experience	58.68 (57.12)	50.61 (54.87)	38.39 (24.91)	33.66 (22.59)	29.37 (17.59)	29.32 (50.62)
After experience	67.83 (72.88)	59.71 (61.84)	50.76 (54.44)	44.73 (4.13)	39.00 (34.53)	28.83 (44.51)
Before round 2		67.63 (76.49)	58.37 (70.81)	49.12 (52.97)	47.05 (52.73)	20.59 (36.17)
Before round 3			41.10 (48.58)	35.27 (42.62)	32.22 (40.17)	8.88 (15.52)
Before round 4				28.76 (23.48)	25.32 (18.86)	3.44 (7.07)
Before round 5					21.71 (12.08)	

NOTE.—Standard deviations are in parentheses.

TABLE 5

RESULTS OF EXPERIMENT 3 (DEBIASING EXPERIMENT)

	Round 1	Round 2	Round 3	Round 4	Slope
Traces completed	8.28 (6.34)	14.48 (6.84)	19.32 (7.11)	24.20 (8.03)	15.92 (4.64)
Prediction:					
Before experience	12.60 (12.01)	15.60 (12.70)	17.88 (13.53)	19.92 (15.00)	7.32 (5.65)
Before experience (debias)	11.48 (10.54)	13.96 (11.18)	16.36 (11.94)	18.20 (13.08)	6.72 (4.61)
After experience	5.32 (6.90)	6.80 (8.92)	8.00 (10.02)	9.36 (12.00)	4.04 (5.22)
After experience (debias)	5.44 (6.97)	6.92 (8.93)	8.12 (10.07)	9.52 (12.03)	4.08 (5.24)
Before round 2		10.00 (6.76)	11.96 (7.78)	13.64 (9.05)	3.64 (2.78)
Before round 3			17.92 (7.84)	20.76 (9.32)	2.84 (1.95)
Before round 4				23.24 (8.30)	. . .

NOTE.—Standard deviations are in parentheses.

Rate-of-Learning Predictions. Before initial experience, participants underpredicted the slope of their own learning curve ($M[A_4 - A_1]$ - $M[P_{BE4} - P_{BE1}] = 8.60$; $F(1, 24) = 30.99$, $p < .001$), and this error increased further after initial experience ($M[A_4 - A_1]$ - $M[P_{AE4} - P_{AE1}] = 11.88$; $F(1, 24) = 66.75$, $p < .001$) and continued before rounds 2 and 3. The results replicated the findings from experiments 1 and 2 and indicated a systematic underconfidence in predicting the rate of learning. Predictions before experience were already pessimistic but became even more so after initial experience. The inaccuracy was greatest after initial experience and diminished in successive rounds.

Results—Debiasing

Before-Experience Predictions in the Debias Condition. As can be seen in table 5, participants responded to the debiasing manipulation by lowering their before-experience predictions, a significant drop ($F(1, 24) = 8.87$, $p < .01$) that made them more accurate, although still overconfident; they predicted that they would complete 11.48 (SD = 10.54) traces but actually completed only 8.28 (SD = 6.34), a significant difference ($F(1, 24) = 2.89$, $p < .05$). Thus, debiasing reduced, but did not eliminate, overconfidence.

After-Experience Predictions in the Debias Condition. The results are consistent with experiment 1; participants were pessimistic after initial experience. Participants predicted that they would complete 5.32 traces (SD = 6.90) after initial experience. After being reminded of the debiasing information (told that most people "predict that they will do worse than they actually do"), participants did increase the prediction to 5.44 traces (SD = 6.97), although the increase in predictions was not significant ($F(1, 24) = .68$, NS). During debriefing, several participants reported that they had not altered their predictions after practicing because they felt they had already incorporated the information about the effect into their first after-experience predictions. After reading the debiasing information, participants remained underconfident in both current ($F(1, 24) = 30.38$, $p < .001$) and rate-of-learning predictions ($F(1, 24) = 69.73$, $p < .001$).

Although providing explicit information about the effect increased prediction accuracy, the correction was insufficient to mitigate the effect. The fact that the effect persisted even after such strong feedback shows it was due not to a lack of information but rather to a failure to use the information. The failure of the debiasing manipulation suggests that directly informing new learners of the pattern of misprediction does not appear to be a viable managerial intervention for reducing or eliminating the bias.

EXPERIMENT 4

The object of experiment 4 is to test whether the findings from the earlier experiments were the result of demand artifacts created by eliciting predictions before experience. Potentially, the underconfidence results could be a result of eliciting the initial overly optimistic predictions. Experiment 4 is designed to eliminate this alternative demand-artifact explanation.

Method and Procedure

Seventy-one students participated in the experiment for extra course credit and performance-based payment. Participants learned to type using the Dvorak format keyboard. The Dvorak keyboard layout claims faster typing and less finger movement than the standard QWERTY keyboard. While these claims are controversial (Liebowitz and Margolis 1990), we selected the Dvorak keyboard because of its lack of familiarity and its commercial availability.

Participants were randomly assigned to one of two conditions (single vs. multiple predictions) in a single-factor, between-subjects design. In the multiple-predictions condition, as in earlier experiments, predictions were elicited before initial experience and after round 1. In the single-prediction condition, participants made performance predictions only after completing round 1 of the task.

The same procedures were followed as in previous experiments. First, participants were given 2 minutes to view the Dvorak keyboard layout and a copy of the words they would be asked to type. Next, participants predicted the number of words they would type. Participants were given 2 minutes to practice with the keyboard. At the end of each

TABLE 6

RESULTS OF EXPERIMENT 4 (KEYBOARD TASK)

	Single prediction		Multiple predictions	
	Round 1	Round 2	Round 1	Round 2
Actual words typed	14.56 (5.71)	20.00 (4.88)	17.14 (7.54)	21.60 (6.65)
Before-experience prediction			32.11 (16.33)	40.80 (20.94)
After round 1 prediction		17.92 (6.03)		19.57 (7.31)

NOTE.—Standard deviations are in parentheses.

round, participants were instructed that the screen would display the gross words per minute, the number of errors typed, and the net words per minute. Finally, as in earlier experiments, the actual task commenced. After round 1, participants who had been assigned to the multiple-predictions condition predicted the number of words they would be able to type. Participants were paid on the basis of their performance ($.03 × number of words typed) or prediction accuracy ($.03 × number of words typed − |number of words typed − number of words predicted|). The average payment to each participant was $1.04.

Results and Discussion

As can be seen in table 6, the results from the previous experiments were replicated. In the condition in which participants made multiple predictions, they were initially overconfident ($M\,[P_{\mathrm{BE1}} - A_1] = 14.97$; SD = 15.30; $F\,(1, 34) = 33.52, p < .001$). After round 1, in both conditions, they significantly underpredicted their performance ($M[P_{22} - A_2]_{\mathrm{single}} = -2.08$; SD = 6.60; $F(1, 34) = 3.29$, $p < .05$; $M[P_{22} - A_2]_{\mathrm{multiple}} = -2.03$; SD = 6.62; $F(1, 35) = 3.59$, $p < .05$). Comparing the predictions made in both conditions, there was no significant difference between the round 2 predictions ($F\,(1, 69) = 1.08$, NS). There is also no significant difference between the round 2 prediction error ($F\,(1, 69) = .001$, NS) when comparing both conditions. Thus, we conclude that the process of eliciting the initial optimistic predictions did not affect subsequent predictions.

This experiment eliminated the alternative explanation that demand artifacts lead to performance underpredictions. The tendency to underpredict persevered, even in the absence of elicitation before experiencing the task. Thus, we conclude that the effect was not a vestige of measurement effects or an overcorrection of early, overly optimistic predictions.

EXPERIMENT 5

Prior research suggests that consumers' valuations of products are dynamic and increase as experience with the product grows (Loewenstein and Strahilevitz 1998). For products that require consumers to acquire skills, product valuations should be related to self-predictions of future performance. Thus, we should expect that for skill-based products, the increased pessimism with experience that has been documented in the prior four studies will lead to a decrease in valuation immediately after initial experience with a product. The objective of this experiment is to investigate whether valuations for skill-based products decline after initial experience.

Method and Procedure

Thirty-three students participated in the experiment for extra credit. We again selected typing with the Dvorak keyboard as the learning task. The same procedures as in experiment 4 were followed, except that there was only one round in the task. Having already demonstrated the drop in confidence using subjective measures in the prior study, in this study we only asked participants to provide their valuation of the Dvorak keyboard. Note that this procedure is conservative in the sense of avoiding the potential demand effect that would have been present if subjects first made subjective judgments and subsequently provided valuations. Participants were asked to value the keyboard before experience and after round 1. The valuations were obtained using a variation of the Becker, DeGroot, and Marshak (1964) procedure. Participants were instructed to state the amount of money that the Dvorak keyboard was worth to them. They were told that after the experiment, 10% of the participant responses were to be randomly selected for inclusion in an actual drawing for the keyboards. For each participant selected, a random number between $0 and $40 was drawn. If the number drawn was less than the participant's valuation of the keyboard, then he or she was given the keyboard. If the number drawn was more than the participant's valuation of the keyboard, then the participant received the number drawn in cash (five actual exchanges were selected, with an average payment of $18.60 to payee recipients—an average of $2.82 per participant).

Results and Discussion

Before the initial experience, the average willingness to pay for the Dvorak keyboard was $8.30 (SD = 10.34). The willingness to pay decreased to $6.58 (SD = 9.16) after experience with the task. This reduction in product valuation after experience with the Dvorak keyboard was significant (M = $1.72; $F(1, 32) = 5.83$, $p < .05$).

The results showed that for products that require acquisition of skills, the pessimism resulting from the effect translated into reduced valuation for the product. We found that valuations for a skill-based product initially decreased,

TABLE 7

RESULTS OF EXPERIMENT 6 (MIRROR-TRACING TASK)

	Round 1	Round 2	Round 3	Round 4	Slope
Traces completed	5.89 (6.37)	9.60 (7.51)	12.89 (8.06)	15.66 (8.80)	9.77 (5.72)
Prediction:					
Before experience	7.87 (7.06)	9.68 (8.15)	11.11 (9.16)	12.06 (9.77)	4.19 (3.98)
After experience	4.49 (5.44)	5.34 (5.72)	6.06 (6.03)	6.87 (6.61)	2.38 (2.10)
Before round 2		7.53 (7.02)	8.28 (7.56)	8.98 (7.93)	1.45 (1.77)
Before round 3			11.53 (9.06)	12.38 (9.46)	.85 (1.12)
Before round 4				14.57 (9.44)	

NOTE.—Standard deviations are in parentheses.

rather than increased, with experience. This decrease in valuation after initial product experience has important managerial implications, suggesting that limited product trials of skill-based products could be detrimental to product sales. Targeting skill-based product sales before customers try the task, while they are optimistic, or perhaps after they have passed through the difficult initial phase would be most successful.

EXPERIMENT 6

We have theorized that the underprediction of learning rates is due to the lack of appreciation of the change from system 2 to system 1 processing. Experiment 6 tests this account. We also examine the role of perceived task difficulty in explaining performance underprediction.

Method and Procedure

Forty-seven subjects participated in the experiment for a $4 show-up fee and performance-based pay (average payment of $14.32 per participant). The mirror-tracing task described in experiment 1 was used after a similar procedure. In addition to predictions, we also obtained measures of the type of processing.

In a review of the literature on dual processing, Evans (2007) identifies two important features of system 1 processing: nonconsciousness and efficiency. Following Menon and Raghubir (2003), we measured nonconsciousness by asking subjects to rate the mental effort, concentration, and thought required to perform the task. Additionally, we asked participants to rate the difficulty of the task. Participants rated each measure on a 7-point scale anchored at not a lot/a lot (for the mental effort/concentration and thought measure) and not at all/very (for the difficulty measure). A nonconscious effort index was formed by averaging the scores for the mental effort/concentration and thought measures (Cronbach's $\alpha = .88$).

With system 1 processing, attention resources become available so that the processing of information that is unrelated to task performance becomes easier and more efficient. Perceptions of efficiency were measured by asking subjects two questions: "Suppose you had to hold a *nine digit number in your mind* (*conversation*) while doing the task. Would that have made it easier or more difficult to do the task?" Participants responded on a 5-point scale anchored at much easier/much more difficult. An efficiency index was formed by averaging the scores on these two questions (Cronbach's $\alpha = .71$). After trial 1, measures of nonconsciousness and efficiency were obtained, and participants were also asked to predict their type of processing at trial 4. At trial 4, the measures were administered again.

Results and Discussion

As in experiments 1–3, the power function ($r^2 = .99$) provided a better fit to the data than did the exponential function ($r^2 = .98$; see table 1). The results of experiments 1–3 were replicated, as can be seen in table 7, and are not reported for brevity. The processing perceptions and predictions are summarized in table 8. It can be observed that participants correctly anticipated a shift from system 2 to

TABLE 8

EXPERIMENT 6 (PROCESSING PERCEPTION)

	After-experience processing perception	Processing prediction for trial 4	After–trial 4 processing perception
Difficulty	5.72 (.17)	4.95 (.20)	4.17 (.24)
Mental effort required	5.38 (.22)	4.89 (.20)	4.23 (.25)
Thought required	5.19 (.23)	4.72 (.22)	4.21 (.26)
Concentration required	6.00 (.16)	5.42 (.19)	4.89 (.22)
Multitask:			
9-digit number	4.23 (.11)	4.11 (.11)	3.96 (.10)
Conversation	4.25 (.11)	4.13 (.12)	4.11 (.10)

NOTE.—Smaller average numbers indicate relatively greater system 1 processing; standard errors are in parentheses.

system 1 processing. Comparing predictions of processing with the actual reported processing measures, however, indicates a failure to appreciate the extent of this shift. Consistent with our expectations, participants significantly underpredicted the extent of the shift as measured by the nonconscious index ($M_{predicted} = 15.04$, $M_{actual} = 13.34$; $F(1, 46) = 7.60$, $p < .01$). The results for the efficiency index ($M_{predicted} = 8.23$, $M_{actual} = 8.06$; $F(1, 46) = .59$, NS) were directionally consistent but not significant.

As in previous experiments, after initially experiencing the task, participants underpredicted ($P_{AE4} - A_4$) peak performance ($M_{P_{AE4}} = 6.87$, $M_{A4} = 15.66$; $F(1, 46) = 73.94$, $p < .001$). To test whether the underprediction of the shift in processing was related to the failure to appreciate the speed of learning effect, we examined the relationship between prediction error in performance and the error in processing predictions. The correlation between the underprediction of performance in trial 4 ($P_{AE4} - A_4$) and the underprediction of processing development (predicted − actual processing measures) revealed a positive relationship as measured by the nonconscious index ($r = .41$, $p < .01$) and the efficiency index ($r = .44$, $p < .01$). These results were consistent with the premise that the effect observed in these experiments is associated with an inadequate appreciation of the speed of the shift in type of processing from system 2 to system 1.

We also examined whether perceived task difficulty during the before-experience learning phase explains the extent of prediction errors. To test the relationship, we computed the correlation between task difficulty and prediction errors for each round. The correlation between task difficulty and the first-round errors ($P_{AE1} - A_1$) was positively significant ($r = .43$, $p < .01$), as well as for the second-round errors ($P_{AE2} - A_2$; $r = .38$, $p < .01$). Correlation was marginally significant for the third-round errors ($P_{AE3} - A_3$; $r = .27$, $p < .06$) but only directional for the fourth-round errors ($P_{AE4} - A_4$; $r = .19$, $p < .25$). Additionally, the correlation between perceived task difficulty and the rate-of-learning predictions ($P_{AE4} - P_{AE1}$) was significant ($r = .39$, $p < .01$). Overall, the results suggest that perception of initial task difficulty explains the extent of errors made in predicting performance.

The results support the explanation that underprediction in performance is related to a failure to appreciate how rapidly system 1 processing overrides system 2 processing. The results also suggest that perceptions of task difficulty play a role in explaining prediction errors.

CONCLUSION

When using skill-based products, consumers must overcome the steep initial learning stage of the skill-acquisition process to fully use all the product's features and benefits. We identify a bias in the early stage of the learning process and propose an explanation for why this learning phase can appear intimidating and, therefore, lead to abandonment of the product.

We have demonstrated consumers' failure to appreciate the speed of learning across tasks involving visual and auditory motor skills, across response modes (units and time), for both skills that were new and those that had to be relearned. We postulate that this type of misprediction is likely to occur for tasks that are novel, require skill development, and for which there is a steep learning curve. Crucially, the effect has behavioral consequences such as lowering valuations of products. Although we eventually observed calibration with experience, we suspect that in many cases such calibration is likely to come too late to prevent initial frustration and attendant behavioral effects. For example, 25% of first-time snowboarders do not take a lesson, purely because of their optimistic expectation that snowboarding will be easy. After trying, 85% of all snowboarders quit and never become long-term participants in the sport (National Ski Areas Association and RRC Associates 2003).

These investigations contribute to an already substantial literature documenting changes in consumers' evaluations of products before and after gaining experience with them. Hamilton and Thompson (2007), for example, show that consumers tend to evaluate products in higher-level, more abstract terms before usage but in more concrete, lower-level terms after gaining experience, which (often belatedly) increases preference for products that are easy to use relative to those that are more desirable but difficult to use. The sudden appreciation of the difficulty of mastering product usage after initial experience that we document in our studies is likely to have a similar effect.

Other studies go beyond documenting changes in post-usage product evaluations and further show that consumers fail to adequately predict such changes. For example, Zauberman (2003) shows that not only do consumers end up getting locked into specific products because of aversion to the small but immediate costs of switching, but they also fail to anticipate such an effect. Meyer et al. (2008) analogously find that consumers are willing to pay higher prices for product features they end up not using. They posit that usage decisions are driven by short-term learning costs that are not taken into account sufficiently at the time of purchase, in part because of overoptimism about use of the new features. Our finding that consumers underpredicted their own learning curves not only is analogous to these misprediction effects but can also exacerbate them because, for example, after gaining initial experience with new product features, consumers are likely to overestimate how long it will take to master them.

Because product valuations are being formed during the initial stage of new product learning, we expect marketing initiatives to be particularly fruitful during this stage of skill-based product adoption (Hoch and Deighton 1989). Our results suggest that the success of new products may be contingent on the degree of learning required. Products, therefore, should be designed so that the skills required are those that the consumer has already mastered or can easily mimic from related tasks. For example, the Nintendo Wii games require players to perform motions that seem natural and familiar, such as punching for boxing, throwing a baseball, and swinging a tennis racket.

Previous literature indicates that consumers attempt to match the capabilities of their purchases to their own perceived skill level (Burson 2007). For example, expert golfers purchase the less forgiving expert-level golf clubs. However, our results suggest that consumers purchasing skill-based products (skis, snowboards, tennis racquets) before experiencing the task will likely select products whose usage difficulty exceeds their capabilities, which would exacerbate their initial difficulties in interacting with the product and accentuate the underestimation of learning effects demonstrated in this article. Consumers who select a skill-based product after initial experience, in contrast, are likely to select products that are easy to use initially but less satisfactory as the consumer gains expertise.

The structure of promotions and incentives to help people persist through the learning curve is also important. For example, promotional schemes targeted at first-time users must be designed so that consumers have sufficient incentives to endure the difficult initial learning phase. Promotions for skill-based products such as a free first lesson, samples, or free trials are likely to be less effective than promotions that help first-time users to achieve a level of expertise at which they are no longer pessimistic about their future learning.

Because of the strong behavioral implications of the effect, firms marketing products or services that require learning should invest resources to hold consumers' hands during the initial stages of product experience or at the least to encourage new customers to persist through the initial phases of learning. For example, one of the authors was involved in the introduction of a national airline's self-check-in kiosk. After initial failure to get consumers to switch from the standard staffed queue, the airline placed service agents at the kiosks to help customers learn the technology.

The findings open up several avenues for further investigation. First, we have examined learning of motor-based skills only. The generalizability of the effect to other tasks and the identification of moderating factors are both worthy of examination. It is plausible, for example, that there are some tasks that are counterintuitively easy (e.g., holding your hand in hot water), in which case we would expect confidence to increase and predictions to become potentially better calibrated after initial experience. Additionally, it could be potentially fruitful to investigate situations in which initial pessimism may not be prevalent (e.g., tasks that are perceived as difficult from the outset). Second, it would be useful to test different methods, such as alternative forms of debiasing, for overcoming the effect. Third, it would be interesting to test alternative explanations for the effect. For example, it is possible that in predicting performance on future trials, people underestimate the slope of skills acquisition because they anchor on their latest performance level and insufficiently adjust for within-trial learning. A better understanding of the psychological underpinnings of the effect could help to explain the effect's persistence, even in the face of explicit hints and despite the ample experiences most people have with having mastered prior tasks.

Much of parenting is about teaching children that persistence pays off—that tasks which initially seem difficult become easier with practice. The results of these studies suggest that, despite whatever lessons our parents might have sought to teach us, most of us have not fully learned the lesson. When the going gets difficult, as tends to happen in the early stages of a task, we tend to become overly pessimistic, making us excessively likely to give up on tasks that could yield long-term fulfillment.

REFERENCES

Ackerman, Phillip L. (1987), "Individual Differences in Skill Learning: An Integration of Psychometric and Information Processing Perspectives," *Psychological Bulletin*, 102 (1), 3–27.

Alba, Joseph W. and J. Wesley Hutchinson (2000), "Knowledge Calibration: What Consumers Know and What They Think They Know," *Journal of Consumer Research*, 27 (September), 123–56.

Anderson, John R. (1982), "Acquisition of Cognitive Skill," *Psychological Review*, 89 (4), 369–406.

Bagozzi, Richard P., Fred D. Davis, and Paul R. Warshaw (1992), "Development and Test of a Theory of Technological Learning and Usage," *Human Relations*, 45 (July), 659–87.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marshak (1964), "Measuring Utility by a Single Response Sequential Method," *Behavioral Science*, 9 (3), 226–32.

Budman, Simon H., David Portnoy, and Albert J. Villapiano (2003), "How to Get Technological Innovation Used in Behavioral Health Care: Build It and They Still Might Not Come," *Psychotherapy: Theory, Research, Practice, Training*, 40 (Spring–Summer), 45–54.

Buehler, Roger, Dale Griffin, and Michael Ross (1994), "Exploring the Planning Fallacy: Why People Underestimate Their Task Completion Times," *Journal of Personality and Social Psychology*, 67 (33), 366–81.

Burson, Katherine A. (2007), "Consumer-Product Skill Matching: The Effects of Difficulty on Relative Self-Assessment and Choice," *Journal of Consumer Research*, 34 (June), 104–10.

Camerer, Colin, George Loewenstein, and Martin Weber (1989), "The Curse of Knowledge in Economic Settings: An Experimental Analysis," *Journal of Political Economy*, 97 (5), 1232–54.

Chein, Jason M. and Walter Schneider (2005), "Neuroimaging Studies of Practice-Related Change: fMRI and Meta-analytic Evidence of a Domain-General Control Network for Learning," *Cognitive Brain Research*, 25 (3), 607–23.

Davis, Fred D. (1989), "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, 13 (September), 319–40.

Davis, Fred D., Richard P. Bagozzi, and Paul R. Warshaw (1989), "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," *Management Science*, 35 (August), 982–1003.

den Ouden, Elke, Lu Yuan, Peter J. M. Sonnemans, and Aarnout C. Brombacher (2006), "Quality and Reliability Problems from a Consumer's Perspective: An Increasing Problem Overlooked by Businesses?" *Quality and Reliability Engineering International*, 22 (July), 821–38.

Dunning, David, Chip Heath, and Jerry M. Suls (2004), "Flawed Self-Assessment: Implications for Health, Education, and the Workplace," *Psychological Sciences in the Public Interest*, 5 (3), 69–106.

Epley, Nicholas and David Dunning (2006), "The Mixed Blessings of Self-Knowledge in Behavioral Prediction: Enhanced Discrimination but Exacerbated Bias," *Personality and Social Psychology Bulletin*, 32 (5), 641–55.

esure Insurance and ICM (2006), "Annual Unused Gadget Poll 2006: New Entrants Fuel Growth of UK's 80m White Gadgets," report, Reigate, http://www.esure.com/news.

Evans, Jonathan (2007), "Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition," *Annual Review of Psychology*, 59 (January), 255–78.

Fischhoff, Baruch (1975), "Hindsight Foresight: The Effect of Outcome Knowledge on Judgment under Uncertainty," *Journal of Experimental Psychology: Human Perception and Performance*, 1, 288–99.

Fitts, Paul M. (1964), "Perceptual Motor Skill Learning," in *Categories of Human Learning*, ed. Arthur W. Melton, London: Academic Press, 243–85.

Fleishman, Edwin A. (1975), "Toward a Taxonomy of Human Performance," *American Psychologist*, 30 (12), 1127–49.

Haier, Richard J., Benjamin V. Siegel Jr., Andrew MacLachlan, Eric Soderling, Stephen Lottenberg, and Monte S. Buchsbaum (1992), "Regional Glucose Metabolic Changes after Learning a Complex Visuospatial/Motor Task: A Positron Emission Tomographic Study," *Brain Research*, 570 (January), 134–43.

Hamilton, Rebecca W. and Deborah V. Thompson (2007), "Is There a Substitute for Direct Experience? Comparing Consumers' Preferences after Direct and Indirect Product Experiences," *Journal of Consumer Research*, 34 (June), 546–55.

Heathcote, Andrew, Scott Brown, and D. J. K. Mewhort (2000), "The Power Law Repealed: The Case for an Exponential Law of Practice," *Psychonomic Bulletin and Review*, 7, 185–207.

Hill, Nicole M. and Walter Schneider (2006), "Brain Changes in the Development of Expertise: Neurological Evidence on Skill-Based Adaptations," in *Cambridge Handbook of Expertise and Expert Performance*, ed. K. Anders Ericsson, Neal Charness, Paul J. Feltovich, and Robert R. Hoffman, New York: Cambridge University Press, 653–82.

Hinds, Pamela (1999), "The Curse of Expertise: The Effects of Expertise and Debiasing Methods on Predictions of Novice Performance," *Journal of Experimental Psychology: Applied*, 5 (2), 205–21.

Hoch, Stephen J. and John Deighton (1989), "Managing What Consumers Learn from Experience," *Journal of Marketing*, 53 (April), 1–20.

Juslin, Peter, Anders Winman, and Henrik Olsson (2000), "Naive Empiricism and Dogmatism in Confidence Research: A Critical Examination of the Hard-Easy Effect," *Psychological Review*, 107 (April), 384–96.

Kahneman, Daniel and Shane Frederick (2005), "A Model of Heuristic Judgment," in *The Cambridge Handbook of Thinking and Reasoning*, ed. K. J. Holyoak and R. G. Morrison, New York: Cambridge University Press, 276–93.

Katz, Michael L. and Carl Shapiro (1986), "Technology Adoption in the Presence of Network Externalities," *Journal of Political Economy*, 94 (August), 822–41.

Koriat, Asher, Hilit Ma'ayan, Limor Sheffer, and Robert A. Bjork (2006), "Exploring a Mnemonic Debiasing Account of the Underconfidence-with-Practice Effect," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32 (3), 595–608.

Koriat, Asher, Limor Sheffer, and Hilit Ma'ayan (2002), "Comparing Objective and Subjective Learning Curves: Judgments of Learning Exhibit Increased Underconfidence with Practice," *Journal of Experimental Psychology: General*, 131 (2), 147–62.

Kruger, Justin (1999), "Lake Wobegon Be Gone! The 'Below-Average-Effect' and the Egocentric Nature of Comparative Ability Judgments," *Journal of Personality and Social Psychology*, 77 (2), 321–32.

Kruger, Justin and Reinhold A. Mueller (2002), "Unskilled, Unaware, or Both? The Better-than-Average Heuristic and Statistical Regression Predict Errors in Estimates of Own Performance," *Journal of Personality and Social Psychology*, 82 (2), 189–92.

Leonard-Barton, Dorothy (1992), *Core Capabilities and Core Rigidities: A Paradox in Managing New Product Development*, New York: Wiley.

Lichtenstein, Sarah, Baruch Fischhoff, and Lawrence D. Phillips (1982), "Calibration of Probabilities: The State of the Art to 1980," in *Judgment under Uncertainty: Heuristics and Biases*, ed. Daniel Kahneman, Paul Slavic, and Amos Tversky, Cambridge: Cambridge University Press, 306–34.

Liebowitz, Stan J. and Stephen E. Margolis (1990), "The Fable of the Keys," *Journal of Law and Economics*, 33 (1), 1–26.

Loewenstein, George (1996), "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Processes*, 65, 272–92.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin (2003), "Projection Bias in Predicting Future Utility," *Quarterly Journal of Economics*, 118, 1209–48.

Loewenstein, George and Michal Strahilevitz (1998), "The Effect of Ownership History on the Valuation of Objects," *Journal of Consumer Research*, 25 (December), 276–89.

Mabe, Paul A. and Stephen G. West (1982), "Validity of Self-Evaluation of Ability: A Review and Meta-analysis," *Journal of Applied Psychology*, 67 (3), 280–96.

McGraw, A. Peter, Barbara A. Mellers, and Ilana Ritov (2004), "The Affective Costs of Overconfidence," *Journal of Behavioral Decision Making*, 17 (4), 281–95.

Menon, Geeta and Priya Raghubir (2003), "Ease of Retrieval as an Automatic Input in Judgment: A Mere Accessibility Framework?" *Journal of Consumer Research*, 30 (2), 230–43.

Meyer, Robert J., Shenghui Zhao, and Jin K. Han (2008), "Biases in Valuation vs. Usage of Innovative Product Features," *Marketing Science*, 27, 1083–96.

Moore, Don A. and Daylian M. Cain (2006), "Overconfidence and Underconfidence: When and Why People Underestimate (and Overestimate) the Competition," *Organizational Behavior and Human Decision Processes*, 103, 197–213.

Moore, Don A. and Tai G. Kim (2003), "Myopic Social Prediction and the Solo Comparison Effect," *Journal of Personality and Social Psychology*, 85, 1121–35.

Morwitz, Vicki G. (1997), "Why Consumers Don't Always Accurately Predict Their Own Future Behavior," *Marketing Letters*, 8 (January), 57–70.

Murray, Kyle B. and Gerald G. Haubl (2007), "Explaining Cognitive Lock-in: The Role of Skill-Based Habits of Use in Consumer Choice," *Journal of Consumer Research*, 34 (June), 77–88.

National Ski Areas Association and RRC Associates (2003), "Model for Growth: Beginner Conversion Research," report, Boulder, CO 80301.

Newell, Allen and Paul S. Rosenbloom (1981), "Mechanisms of Skill Acquisition and the Law of Practice," in *Cognitive Skills and Their Acquisition*, ed. John R. Anderson, Hillsdale, NJ: Erlbaum, 1–55.

Rogers, Everett M. (1976), "New Product Adoption and Diffusion," *Journal of Consumer Research*, 2 (March), 290–301.

Ryan, Bryce and Neal C. Gross (1943), "The Diffusion of Hybrid Seed Corn in Two Iowa Communities," *Rural Sociology*, 8 (January), 15–24.

Schneider, Walter and Jason M. Chein (2003), "Controlled and Automatic Processing: Behavior, Theory, and Biological Mechanisms," *Cognitive Science*, 27 (3), 525–59.

Snoddy, George (1926), "Learning and Stability," *Journal of Applied Psychology*, 10 (February), 1–6.

Stanovich, Keith and Richard West (2002), "Individual Differences in Reasoning: Implications for the Rationality Debate?" in *Heuristics and Biases: The Psychology of Intuitive Judgment*, ed. T. Gilovich, D. Griffin, and D. Kahneman, New York: Cambridge University Press, 421–40.

Thompson, Deborah V., Rebecca W. Hamilton, and Roland T. Rust (2005), "Feature Fatigue: When Product Capabilities Become Too Much of a Good Thing," *Journal of Marketing Research*, 42 (November), 431–42.

Tirole, Jean (1988), *Theory of Industrial Organization*, Cambridge, MA: MIT Press.

Venkatesh, Viswanath and Fred D. Davis (2000), "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies," *Management Science*, 46 (February), 186–204.

Walker, Dave (2007), "1.25 Billion Wasted on Gadgets Due to Fear and Ignorance," *Techdigest*, May 2, http://techdigest.tv/2007/05/125bn_wasted_on.html.

Zauberman, Gal (2003), "The Intertemporal Dynamics of Consumer Lock-in," *Journal of Consumer Research*, 30 (December), 405–19.