



## The effort heuristic

Justin Kruger,<sup>a,\*</sup> Derrick Wirtz,<sup>a</sup> Leaf Van Boven,<sup>b</sup> and T. William Altermatt<sup>c</sup>

<sup>a</sup> Department of Psychology, University of Illinois at Urbana-Champaign, 603 E. Daniel Street, Champaign, IL 61820, USA

<sup>b</sup> University of Colorado at Boulder, Boulder, CO, USA

<sup>c</sup> University of Michigan-Flint, Flint, MI, USA

Received 7 November 2002; revised 17 March 2003

### Abstract

The research presented here suggests that effort is used as a heuristic for quality. Participants rating a poem (Experiment 1), a painting (Experiment 2), or a suit of armor (Experiment 3) provided higher ratings of quality, value, and liking for the work the more time and effort they thought it took to produce. Experiment 3 showed that the use of the effort heuristic, as with all heuristics, is moderated by ambiguity: Participants were more influenced by effort when the quality of the object being evaluated was difficult to ascertain. Discussion centers on the implications of the effort heuristic for everyday judgment and decision-making.

© 2003 Elsevier Science (USA). All rights reserved.

*Keywords:* Judgment and decision-making; Heuristics and biases; Social cognition; Effort

Consider the following: After receiving a low grade on a paper, a student approaches you after class and complains about his or her grade. Occasionally, such protests stem from errors in grading or strategic attempts by the student to “work” the system. But more often, the grade is correct, and the protest sincere. In such cases, it is not uncommon to hear some version of the following: “But I don’t understand. I worked so hard...how could I get a D?” Apparently, the student believes that because he or she invested great effort, the outcome of that effort must be great as well. Indeed, we recently asked a group of Cornell University students to indicate (confidentially) how much effort they put into a term paper and the grade they expected to receive, and compared those estimates with the actual grades received. Although predicted and actual grades were positively correlated, that relationship was dwarfed by the relationship between the grade the students expected to receive and the effort they reported investing in the paper.

This result will perhaps come as no surprise to social psychologists. The notion that effort can influence judgment is an old and important one in social psychology. Some of the earliest work in the dissonance and self-perception paradigms, for instance, demonstrated that the more effort individuals invest—whether in the form of time, physical exertion, pain, or money—the more positively they evaluate the product of that effort (Aronson & Mills, 1959; Bem, 1972; Festinger, 1957; Gerrard & Mathewson, 1966; Wicklund & Brehm, 1976). For instance, the “hazing” of fraternity pledges can make the goal of fraternity membership seem more attractive, in part because the inconsistency between the tremendous effort invested and the moderate value of the goal creates dissonance—dissonance that can be reduced by inflating one’s evaluation of club membership (Aronson & Mills, 1959; Gerrard & Mathewson, 1966). Following this logic, it is possible that the student in the opening example believed that he or she deserved a better grade than the one received because of the dissonance created by spending hours on a D paper.

We suspect, however, that there is an additional route by which effort influences judgment, one distinct from both dissonance theory and self-perception theory in

\* Corresponding author.

E-mail addresses: [jkruger@uiuc.edu](mailto:jkruger@uiuc.edu) (J. Kruger), [d-wirtz@uiuc.edu](mailto:d-wirtz@uiuc.edu) (D. Wirtz), [vanboven@colorado.edu](mailto:vanboven@colorado.edu) (L. Van Boven), [altermat@um-flint.edu](mailto:altermat@um-flint.edu) (T.W. Altermatt).

both explanation and application. We propose that people use effort as a heuristic for quality. The more effort invested in an object—be it a painting, a poem, or a paper submitted to *JESP*—the better it is deemed to be. To be sure, there are exceptions, but all else equal people tend to believe that a painting that takes 2 days to paint is better than one that takes 2 h, just as people tend to believe that an additional manuscript revision will result in a better paper. If so, then the student in the opening example may have believed that he or she deserved a better grade because the time and effort invested was used as a heuristic for the quality.

As an analogy, consider the use of availability as a heuristic for frequency described by Daniel Kahneman and Amos Tversky (Schwarz et al., 1991; Tversky & Kahneman, 1973). These researchers found that the ease with which specific instances of a category come to mind is used to estimate the frequency of instances of that category. For instance, people believe (correctly) that more words begin with the letter *r* than the letter *z* because it is easier to think of words that begin with the letter *r* than the letter *z*.

The reason people use the availability heuristic is that, first, frequency can be difficult to assess, and second, the heuristic is ecologically valid (although this latter point is often overlooked). Accurately judging number of instances of a category—be it the number of words that begin with the letter *r* or the number of times one has washed the dishes in a given month—is a formidable task. And the ease with which instances of a category can be brought to mind is generally a reliable indicator of the frequency of instances of that category. Availability tells us, quite correctly, that there are more words that begin with the letter *r* than the letter *z*, just as it tells us, also quite correctly, that we have washed the dishes more often in the past month than we have cleaned the toilet.

In much the same way, we argue, and for much the same reason, people use effort as a heuristic for quality. Like frequency, quality can be difficult to determine: The monetary value of a painting, the quality of a committee decision, the theoretical contribution of a manuscript—all can be difficult qualities to assess. And just as availability is for frequency, effort is a generally reliable indicator of quality. All else equal, paintings that receive the lion's share of an artist's attention tend to be superior to those that receive less of her attention, just as spending more time revising a manuscript usually results in a better article.

But note that the association between effort and quality is imperfect. As writers know, although hard work and careful revision are the hallmarks of good prose, a quick moment of inspiration can occasionally rival what would otherwise take hours to produce. Indeed, it is not uncommon for one afternoon of writing to produce several pages of useful text, and

another to pass by with only a paragraph or two to show for it (we call the latter “Fridays”). The same is true of virtually all creative domains. Thus, although effort is generally a valid cue to quality, the effort heuristic, as with all heuristics, can occasionally lead to error (cf., Chaiken, Liberman, & Eagly, 1989; Gilovich, Griffin, & Kahneman, 2002; Kahneman & Frederick, 2002; Kelley & Jacoby, 1998; Schwarz, 2002; Slovic, Finucane, Peters, & MacGregor, 2002; Tversky & Kahneman, 1974).

The present research was designed to test the hypothesis that people use effort as a cue for quality, and this can lead to error. Our strategy for demonstrating the former was by demonstrating the latter. We conducted 3 experiments in which participants made judgments of quality—of a poem in Experiment 1, paintings in Experiment 2, and medieval arms and armor in Experiment 3. In each, we manipulated the effort ostensibly invested in the object's creation. Despite the fact that the actual quality of the work remained the same (by virtue of the fact that all participants in a given study rated the same object or objects), we expected the effort manipulation to influence perceived quality: the greater the effort, the greater the perceived quality.

A key distinction between the present and past research on the role of effort on judgment, and work in the dissonance paradigm in particular, is that we focused on other-generated effort rather than self-generated effort. This allowed us to examine the influence of effort on judgment in a manner immune to a dissonance explanation: Whereas dissonance applies to judgments about the fruits of one's own labors, it does not apply to judgments about the efforts of others. The effort heuristic, in contrast, implies that effort is used as a heuristic for quality even when someone else produces the work in question.

We also examined a potential moderator of the use of the effort heuristic: ambiguity. If effort is a heuristic for quality, then it ought to be utilized more when quality is otherwise difficult to ascertain—as is the case, we argue, in many artistic domains—and less (though perhaps still to some extent) when evaluating an object that is less ambiguous in terms of quality, such as a television set (whose picture quality is readily apparent) or a 3-week old filled of sole. Thus, in Experiment 3, in addition to manipulating the effort ostensibly invested by the product's creator, we also manipulated the ambiguity of the quality of the object being evaluated.

### Experiment 1: Poetry

Participants read and evaluated the poem “Order” by contemporary poet Michael Van Walleghen. Half were told that it took Van Walleghen 4 h to compose the poem (low-effort condition), and half were told that it

took 18 h (high-effort condition). We predicted that participants would evaluate the poem more favorably when told that it took 18 h to complete than when told it took 4 h.

### Method

One hundred and forty-four Williams College students enrolled in an introductory course in psychology earned course credit for participating. On arrival to the lab, participants were told that the experiment concerned the way in which people evaluate of poetry. Participants read “Order” by Van Walleghen (1966). Next to the poem was a list of details about the poem—some genuine and some not—including the title, author, age of poet, and the time he ostensibly spent writing the poem. This information was constant across participants with the exception of the time spent by the poet: participants randomly assigned to the low-effort condition were told that Van Walleghen spent 4 h on the poem, whereas participants in the high-effort condition were told that the poet spent 18 h. The non-effort information was included because pretesting revealed that this reduced awareness of time as the independent variable. This, along with the between-participant nature of the design, minimized potential demand characteristics.

After reading the poem, participants recorded how long it took to compose, as well as each additional piece of information provided about the poem (the latter to further reduce demand). Next, they evaluated it. First, they indicated how much they liked the poem on a scale from 1 (*hate it*) to 6 (*it's OK*) to 11 (*love it*). Then, they assessed the “overall quality” of the poem on a scale from 1 (*terrible*) to 6 (*OK*) to 11 (*excellent*). Finally, participants indicated how much money the poem would be likely to fetch (in US dollars) if sold to a poetry magazine.

### Results and discussion

Four participants failed to complete the dependent measures, and two participants were extreme outliers with responses that exceeded 5 standard deviations from the mean. These participants were excluded from the analysis, yielding a final sample of 138 (89 women, 45 men, and 4 unidentified). Gender did not influence any of the results in this or any of the experiments in this research and will not be discussed further.

Participants' responses to the two Likert-scale questions were averaged to create a single index ( $r = .84$ ,  $p < .001$ ). As Table 1 shows, participants provided more favorable evaluations of the poem when they thought it took Walleghen 18 h to compose than when they thought it took him 4 h,  $F(1, 136) = 3.98$ ,  $p = .048$ , and also thought the poem was worth more money,

Table 1  
Mean evaluation of Van Walleghen's “Order” by effort condition, Experiment 1

	Effort condition	
	Low effort	High effort
Mean liking/overall quality	5.84	6.43
Median perceived value	\$50	\$95

$F(1, 134) = 3.62$ ,  $p = .059$  (after a log transformation to satisfy the ANOVA assumption of normality).<sup>1</sup>

These results provide initial evidence that effort is used as a heuristic for evaluations of quality. This was true both for participants' subjective evaluations of the poem and for their estimates of the objective value of the poem—i.e., the poem's monetary value.

### Experiment 2: Painting

Our next experiment was designed to extend the results of the previous experiment to a different artistic domain (painting), and to see whether effort would be used as a heuristic for quality even among people with self-proclaimed expertise in the subject at hand. In keeping with previous work (e.g., Tversky & Kahneman, 1971), we reasoned that the use of effort among self-perceived experts would constitute a conservative test of the heuristic.

An additional change from the previous experiment was from a fully between-participant design to a mixed-model design. Participants evaluated the quality of two paintings rather than one, with the effort ostensibly invested by the artist varying not only between-participants but also within. Specifically, participants evaluated (in random order) 2 paintings by Deborah Kleven, a contemporary fine and graphic artist based in Washington, DC: “12 Lines” and “Big Abstract.” Half of the participants were told that the former took Kleven 4 h to paint and the latter 26 h, and the other half were told the opposite. After rating each painting separately, participants compared the two paintings directly. This enabled us to examine whether effort influences not only absolute judgments of quality, but also *comparative* judgments, to see whether participants' preferences for one painting over the other could be changed by varying the effort ostensibly invested by the artist.

### Method

**Participants.** Non-experts (21 women, 12 men)—primarily undergraduate psychology majors—were recruited from an introductory psychology course at

<sup>1</sup> The degrees of freedom vary slightly from analysis to analysis due to incomplete data.

Cornell University and earned \$6 for their participation. Self-identified experts (23 women, 10 men)—primarily undergraduate or graduate art students—were recruited via an “art experts wanted” advertisement posted at several locations on the Cornell University campus and were paid \$10 for their participation.

**Procedure.** Participants were recruited individually for an experiment of “artistic judgment.” Participants were shown 2 paintings by Deborah Kleven: “12 Lines” and “Big Abstract.” Next to each painting was a list of several details about the painting, including the name and nationality of the artist (Deborah Kleven, American), the title of the piece (12 Lines or Big Abstract), the media used (oil and canvas or oil and enamel on canvas), the age of the artist upon completion of the painting (37 or 49), and, of key importance, the “total time spent by artist on painting.” Half of the participants were told that 12 Lines took Kleven 4 h to paint and Big Abstract took 26 h, and the other half were told the opposite. The effort information was orthogonal to all of the other information presented, and the order in which the two paintings were evaluated was counter-balanced across participants.

Participants first rated each painting separately. As in Experiment 1, participants indicated how much they liked the painting and the overall quality of the painting, each on 1–11 scales. In addition, participants estimated how their fellow classmates would evaluate the overall quality of the painting, again on a 1–11 scale. Finally, participants estimated the amount of money (in US dollars) the painting would probably receive if sold at auction.

After rating each painting individually, participants then compared the two directly. Participants indicated which painting they liked more, which painting is superior in overall quality, which painting a group of their peers would probably think is superior in overall quality, and which painting would fetch more money at auction. In addition to the four dichotomous measures, participants also compared the overall quality of, and effort invested in, each painting on a continuous scale from 1 (painting #1 much more) to 6 (exactly the same) to 11 (painting #2 much more). Finally, participants indicated whether they had seen either of the paintings before.

## Results

The order in which the paintings were presented did not significantly influence any of the results in this experiment and will not be discussed further. No participant reported having seen either of the paintings before.

**Indirect comparisons.** Our primary prediction was that participants’ evaluations of the painting would vary as a function of the perceived time and effort invested by the artist. To find out whether this was the case, we first

averaged the ratings of liking, overall quality, and anticipated peer ratings into composite measures for the two paintings ( $\alpha > .78$ ). We then submitted these indices to a 2 (painting rated: 12 Lines vs. Big Abstract)  $\times$  2 (high-effort painting: 12 Lines vs. Big Abstract)  $\times$  2 (sample: self-identified experts vs. novices) mixed-model ANOVA.

This analysis yielded the anticipated two-way interaction between painting and effort,  $F(1, 62) = 5.99$ ,  $p = .017$ . As Fig. 1 reveals, participants preferred 12 Lines over Big Abstract when they thought 12 Lines took longer to paint, but the opposite tended to be true when they thought that Big Abstract took longer to paint. Importantly, no other main effects or interactions were significant, all  $F_s < 1.3$ . These data indicate that the effect of perceived effort on perceived quality was independent of whether participants had self-professed expertise in the domain.

The effort manipulation had a similar effect on participants’ estimates of how much the paintings were worth. After a log transformation, a 2 (painting rated: 12 Lines vs. Big Abstract)  $\times$  2 (high-effort painting: 12 Lines vs. Big Abstract)  $\times$  2 (sample: experts vs. novices) mixed-model ANOVA yielded three effects. On average, participants thought that Big Abstract was worth more money than 12 Lines,  $F(1, 59) = 5.51$ ,  $p = .022$ , and self-identified art experts tended to provide higher estimates than novices,  $F(1, 59) = 2.92$ ,  $p = .092$ . Importantly, this analysis also revealed the expected two-way interaction between painting and effort,  $F(1, 59) = 10.21$ ,  $p = .002$ . As Table 2 reveals, participants who thought 12 Lines took longer to produce thought that it

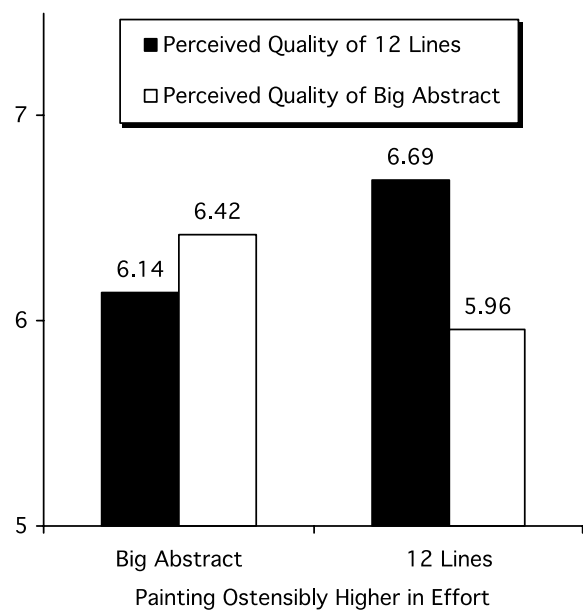


Fig. 1. Perceived quality of Deborah Kleven’s 12 Lines and Big Abstract by effort condition, Experiment 2.

Table 2  
Median perceived value of Deborah Kleven's 12 Lines and Big Abstract by effort condition, Experiment 2

	Painting higher in perceived effort	
	12 Lines	Big Abstract
Perceived value of 12 Lines	\$1000	\$1000
Perceived value of Big Abstract	\$650	\$1500

was worth more money than Big Abstract, whereas the opposite tended to be true when participants thought that Big Abstract took longer to paint. No other main effects or interactions were significant.

*Direct comparisons.* The effect of perceived effort on participants' preferences also was evident when participants compared the two paintings directly. As Table 3 shows, participants were more likely to prefer 12 Lines over Big Abstract when they thought the former took the artist longer to paint than when they thought the converse. This interaction was significant for all four dependent measures,  $\chi^2$ s > 4.1,  $ps < .05$ . Responses to the continuous measure revealed a similar pattern: Participants preferred 12 Lines over Big Abstract to a greater extent when they thought the former took longer to produce ( $M = 7.14$ ) than when they thought the latter took longer to produce ( $M = 6.00$ ),  $F(1, 64) = 5.26$ ,  $p = .025$ .

*Mediational analysis.* Recall that in addition to comparing the two paintings in terms of quality, participants also compared them in terms of the effort invested by the artist. This enables a test of the extent to which the influence of the time manipulation on perceived quality is mediated by perceived effort. We examined this relationship by using a structural equation modeling program within the AMOS procedure (Arbuckle & Wothke, 1999). As Fig. 2 shows, whereas the time manipulation was significantly related to perceived effort,  $Z = 5.76$ ,  $p < .001$ , and perceived effort significantly related to perceived quality,  $Z = 5.32$ ,  $p < .001$ , there was no direct relationship between time and quality when effort was held constant (nor was there any relationship between self-professed art expertise on any of these variables). Furthermore, a Sobel (1982) test

Table 3  
Proportion of sample preferring 12 Lines over Big Abstract by effort condition, Experiment 2

	Painting higher in perceived effort	
	12 Lines (%)	Big Abstract (%)
Higher in overall quality	67	24
Liked more	67	42
Higher in predicted peer quality rating	53	39
Higher perceived worth	73	27

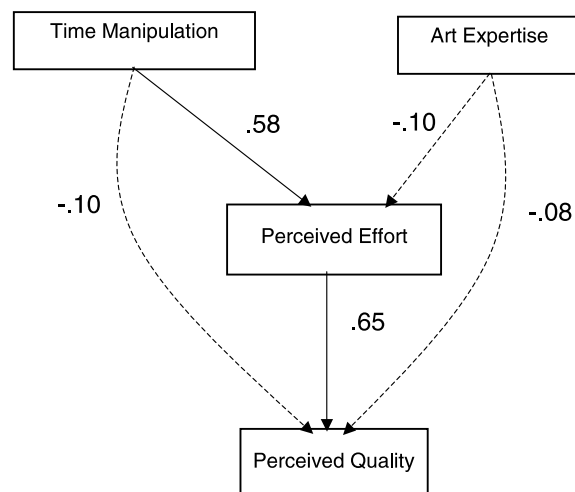


Fig. 2. Relationship between time, self-rated art expertise, perceived effort, and judgments of quality, Experiment 2.

confirmed that the association between time and perceived quality was mediated by perceived effort,  $Z = 2.31$ ,  $p = .021$ .

### Discussion

The results of Experiment 2 bolster our claim that effort is used as a heuristic for quality. As in Experiment 1, the more time invested in the object, the more favorably it was rated. Structural equation modeling confirmed that the effect of the time manipulation on perceived quality was fully mediated by perceived effort.

Interestingly, self-identified art experts did not appear to rely on effort any less (or more) than novices, despite the fact that the former were presumably more practiced at evaluating art, and might be expected to evaluate it based on features inherent in the artistic product. Tversky and Kahneman (1971) interpreted experts' use of the representativeness heuristic as evidence regarding its ubiquity. Analogously, we believe use of effort in judgments of artistic quality among self-identified art experts points to the generality and intuitive appeal of effort as a heuristic for quality.

### Experiment 3: Arms and armor

Our third and final experiment was designed to extend the results of the previous two experiments to a third domain, and also to examine a potential moderator in the use of the effort heuristic: ambiguity. If effort is a heuristic for quality, it follows that it ought to be particularly utilized in judgments of particularly ambiguous stimuli. Thus, in addition to varying the perceived effort invested by the artist, we also varied the ambiguity of the stimulus.

Specifically, participants in Experiment 3 rated the quality of several images of medieval arms and armor presented on a computer screen. When rating the final target piece of armor, half of the participants were told that it took the blacksmith 110 h to complete, and half were told that it took 15 h. The ambiguity of the piece was manipulated by simply altering the resolution of the image: half of the participants viewed a high-resolution image of the piece, and half viewed a low-resolution image. We predicted that the greater the perceived effort, the greater the perceived quality—particularly in the case of the ambiguous, low-resolution image of the object.

### Method

**Participants.** Two hundred and thirty-five University of Illinois undergraduates (133 women, 99 men, and 3 unidentified) participated in the web-based study on a volunteer basis. One of several experimenters recruited participants door-to-door as part of a class project, and only participants who owned a computer with web access were included in the study.

**Procedure.** Once participants agreed to participate in the study they were told that the experiment involved viewing several images of medieval arms and armor and rating each one on a questionnaire. Each participant then accessed the experiment webpage on his or her computer. The webpage sequentially presented 14 images of varying resolution of objects such as a 16th century Spanish morion and a suit of armor etched in the style of the Court of the Dukes of Brunswick. Next to each image was a description of the piece, including the item name, the materials used (e.g., iron with gold and leather), the weight of the object (in pounds), and, importantly, the manufacturing time (in hours).

Participants were instructed to examine the piece of arms or armor, and then rate it on a questionnaire (after first recording the item name, materials, weight, and manufacturing time, the former to reduce suspicion of the true independent variable). The questionnaire presented five 9-point semantic differential scales with endpoints labeled *poorly-crafted/well-crafted*, *cheap/expensive*, *fit for serf/fit for king*, *low quality/high quality*, and *worthless/priceless*, respectively. After participants completed their ratings, this process repeated until every piece was viewed and rated.

The key piece in the study was the last image presented, a late 16th century iron breastplate housed in the Tower of London. Half of the participants were told that the piece took 110 h to complete, and the other half were told that it took 15 h. Participants randomly assigned to the low-ambiguity condition viewed a clear image of the piece, whereas participants assigned to the high-ambiguity condition viewed a blurrier image. The design of the experiment was thus a 2 (object ambiguity:

high vs. low)  $\times$  2 (effort: high vs. low) fully between-participant factorial.

### Results and discussion

Our primary prediction was that the higher the perceived effort invested by the blacksmith, the greater the perceptions of quality—particularly when the object in question was ambiguous. To find out whether this was the case, we first averaged participants' five evaluations into a composite measure ( $\alpha = .96$ ). We then submitted this index to a 2 (object ambiguity: high vs. low)  $\times$  2 (effort: high vs. low) ANOVA, which revealed two main effects. First, participants provided higher ratings when they viewed a clear image of the piece than when they viewed a blurry image,  $F(1, 231) = 57.35$ ,  $p < .001$ . Second, participants provided higher ratings of the piece when they thought it took the blacksmith 110 h to produce than when they thought it took 15 h,  $F(1, 231) = 23.07$ ,  $p < .001$ , consistent with the results of the previous two studies. Of key importance, this analysis also revealed the predicted interaction,  $F(1, 231) = 7.88$ ,  $p = .005$ . As Fig. 3 reveals, the influence of effort on judgment was bigger in the high-ambiguity condition than in the low-ambiguity condition, consistent with our predictions.

We attribute this interaction to the fact that the quality of the armor is more ambiguous in the low-resolution condition than in the high-resolution condition. Participants who viewed the lower-resolution version of the piece had less information about the piece upon which to base their evaluations than participants who viewed a higher-resolution picture. The fine etching, for instance, apparent in the high-resolution version of the image was

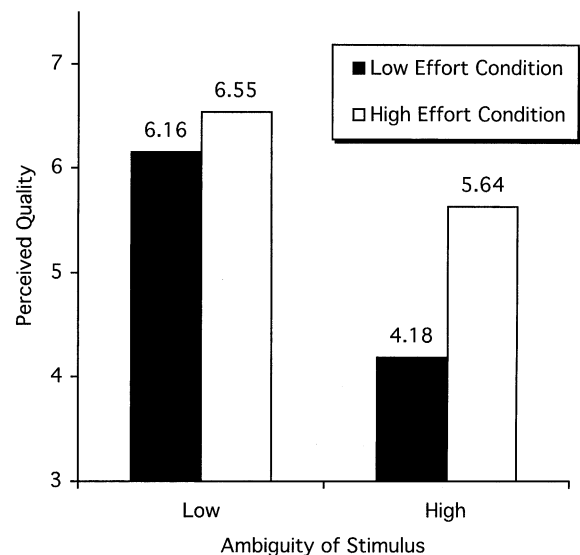


Fig. 3. Evaluations of armor by effort and ambiguity condition, Experiment 3.

virtually invisible in the low-resolution image. As such, participants in the low-resolution condition had less objective information upon which to make a judgment of quality, and thus were particularly likely to rely on the effort ostensibly invested by the blacksmith. Indeed, although participants in both conditions tended to provide more favorable ratings in the high vs. low effort conditions, it was only in the low-resolution condition that the simple effect was significant at the .05  $\alpha$ ,  $F(1, 112) = 26.83$ ,  $p < .001$  vs.  $F(1, 119) = 2.15$ ,  $p = .145$ .

## General discussion

In the late 1940s, Jackson Pollock burst on the international art scene—then dominated by Picasso, Dali, and Magritte—and quickly became one of the “greatest American painter[s] of the 20th Century” (Seiberling, 1949, p. 42). His unconventional “drip method” yielded some of the most prized—and expensive—artworks produced in the last century.

As revolutionary as his technique and paintings are today, many of his contemporary critics (and much of the general public) were unimpressed. “Jackson Pollock’s paintings represent absolutely nothing,” wrote one critic, “chaos, absolute lack of harmony, complete lack of structural organization, total absence of technique” (O’Connor, 1967, pp. 54–55). Others agreed, calling his work “as unpalatable as yesterday’s macaroni,” or asking simply, “is this a painting?” (Friedman, 1972; Seiberling, 1949). Another, in a 1948 “critics roundtable” sponsored by Life Magazine, expressed what was perhaps the most common reaction: “I could have made [it] myself” (Davenport, 1948, p. 62).

What these early critics and laypeople did not know was that they probably could *not* have made the paintings themselves. What many assumed was the result of thoughtless splattering of paint onto a canvas was in fact a slow, deliberate, and exhausting process often requiring weeks, even months of hard work (Goodnough, 1951). What is striking about this assertion is how important it was in establishing Pollock as a legitimate artist. Indeed, the effort invested by Pollock in his work was occasionally used to defend it. As the noted art historian Leo Steinberg put it, “Questions as to the validity of Pollock’s work, though they remain perfectly good in theory, are simply blasted out of relevance by these manifestations of Herculean effort, this evidence of mortal struggle between the man and his art” (cited in Varnedoe & Karmel, 1998, p. 17).

The research presented here suggests that such effort-based judgments of quality are common, and reflect a more general judgmental tendency we term the effort heuristic. Whether evaluating paintings (Experiment 2), poems (Experiment 1), or armor (Experiment 3), the greater the perceived effort invested by the artist, the

better they were assumed to be. This was true not only in terms of subjective liking for the work, but also assessments of objective quality. These results dovetail with those in the dissonance and self-perception literatures showing that the more effort an individual invests in something, the better they assume it to be. However, whereas this previous work focused on self-generated effort, the present research suggests that effort is used to inform judgments of the fruits of others’ labors.

How does the present work fit with previous work on judgmental heuristics? Heuristics range on a continuum from broad judgmental strategies that apply to a wide range of judgments, such as availability as a heuristic for frequency, to narrow judgmental cues that arise in specific contexts, such as height as a heuristic for weight (Gilovich & Griffin, 2002; Kahneman & Frederick, 2002). We believe the effort heuristic falls closer to the broad, general-purpose end of the continuum rather than the narrow, special-purpose end. Everyday life, after all, is replete with judgments of quality. How much money should I pay for a garage sale “treasure?” How favorably should I review a manuscript? What grade should I give a student’s essay? Whether as consumers, reviewers, or teachers, we are often called upon to assess the quality of other people’s creative endeavors. In each case, we suspect that all else equal, the more effort invested in the object the better it is assumed to be.

That said, there are clear caveats to the effort heuristic worth considering. Although effort was consistently associated with increased perceived quality in the present experiments, we suspect that occasionally the opposite may be true. A labor-intensive manuscript may be seen as inferior to one that comes more easily if, for instance, the latter is thought to be a product of inspiration and the former a struggle with logical leaps, overly complicated language, and unruly data. Thus, the *type* of effort involved may be a key moderator, one worthy of future research.

We suspect, however, that the net effect of increased effort on perceptions is more often positive than negative, and that the results presented here are the rule rather than the exception. We are reminded of the advice a journal editor we know includes in his cover letter to reviewers of revised manuscript submissions: “Please remember that our evaluation should rest on whether the revised paper is above the high *JPS* threshold now, and not on how hard the author worked in making the revisions” (E. Diener, personal communication, October 17, 2001). In light of the data presented here, we suspect that this well-reasoned advice may be easier to give than to follow.

## Acknowledgments

This research was supported financially by Research Grant 1-2-69853 from the University of Illinois Board of

Trustees awarded to Justin Kruger. Portions of this research were presented at Society for Personality and Social Psychology 2001 and Midwestern Psychological Association 2002 annual meetings.

## References

- Arbuckle, J. L., & Wothke, W. (1999). *AMOS [Computer software]*. Chicago: SmallWaters Corporation.
- Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology*, *59*, 177–181.
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York: Academic Press.
- Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic processing within and beyond the persuasion context. In J. S. Uleman, & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.
- Davenport, R. W. (1948). A *Life* roundtable on modern art. *Life*, *25*(2), 62.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row & Peterson.
- Friedman, B. H. (1972). *Jackson Pollock: Energy made visible*. New York: McGraw-Hill.
- Gerrard, H. B., & Mathewson, G. C. (1966). The effects of severity of initiation on liking for a group: A replication. *Journal of Experimental Social Psychology*, *2*, 278–287.
- Gilovich, T., & Griffin, D. (2002). Introduction—heuristics and biases: then and now. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 1–18). New York: Cambridge University Press.
- Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. New York: Cambridge University Press.
- Goodnough, R. (1951). Pollock paints a picture. *Artnews*, *50*, 38–41, 60–61.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). New York: Cambridge University Press.
- Kelley, C. M., & Jacoby, L. L. (1998). Subjective reports and process dissociation: Fluency, knowing, and feelings. *Acta Psychologica*, *98*, 127–140.
- O'Connor, F. V. (1967). *Jackson Pollock*. New York: The Museum of Modern Art.
- Schwarz, N. (2002). Feelings as information: Moods influence judgments and processing strategies. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 534–547). New York: Cambridge University Press.
- Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, *61*, 195–202.
- Seiberling, D. (1949). Jackson Pollock: Is he the greatest living painter in the United States? *Life*, *27*, 42–45.
- Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equations models. In S. Leinhardt (Ed.), *Sociological methodology* (pp. 290–312). San Francisco: Jossey-Bass.
- Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 397–420). New York: Cambridge University Press.
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*, 105–110.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.
- Van Walleghen, M. (1966). *The Wichita Poems*. Urbana, IL: University of Illinois Press.
- Varnedoe, K., & Karmel, P. (1998). *Jackson Pollock*. New York: The Museum of Modern Art.
- Wicklund, R., & Brehm, J. (1976). *Perspectives on cognitive dissonance*. Potomac, MD: Erlbaum.