

Instructor Fluency Correlates with Students' Ratings of Their Learning and Their Instructor in an Actual Course

Michael J. Serra, Debbie A. Magreehan

Department of Psychological Sciences, Texas Tech University, Lubbock, TX, USA

Email: michael.serra@ttu.edu

Received 28 May 2016; accepted 31 May 2016; published 3 June 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The experience of ease or fluency that occurs when learners acquire information is often highly related to their metacognitive judgments of learning for that information. Laboratory-based research indicates that fluency can contribute to students' overconfident judgments of learning and predictions of future test performance. Such research, however, typically involves artificial learning situations presented for brief periods of time and without a strong investment on the part of the learners. In actual courses, the most likely source of fluency may be instructor fluency: the experience of fluency that stems from content-independent attributes of the instructor and his or her presentation of the information. To examine whether this form of fluency relates to students' judgments of learning in actual academic courses, we include a measure of instructor fluency in a survey completed by college students ($n = 606$) answering questions about their course instructors. Students' content-independent perceptions of instructor fluency (e.g., volume; eye contact) are related to their judgments of learning for the course content, to their ratings of various qualities of the instructor and the course, and to their self-reported interest and motivation in the course. Importantly, these relationships maintain when we control for students' final grades in the course and despite the fact that students make these ratings at a long temporal delay from the classroom experience. Therefore, much as occurs in the laboratory, students' metacognitive judgments of their learning and ratings of instructor attributes are related to content independent qualities of their course instructors in actual semester-long courses.

Keywords

Fluency, Judgments of Learning, Judgment Bias, Student Ratings, Teaching Evaluations

1. Introduction

From teaching (Williams & Ceci, 1997) and research presentations (Naftulin, Ware Jr., & Donnelly, 1973) to corporate proposals (Duarte, 2012) and political debates (Jackson-Beeck & Meadow, 1979), people's evaluations of the quality of speakers and their ideas are greatly influenced by various attributes of the presenter and their presentation. These attributes include content-specific qualities of the presenter or the presentation itself (e.g., the quality of information and arguments presented), but also include content-independent qualities of the presenter or the presentation itself (e.g., the speaker's expressiveness or body language; the design of Power-Point presentations). These latter qualities can contribute to the experience of a "fluent" (easily-processed) experience on the part of audience. Importantly, audience members often interpret the experience of fluency as a heuristic indicative of quality. Put differently, the audience may rate the quality of a presentation as being higher when it is experienced fluently compared to if it is experienced in a disfluent way, even if the content is the same.

In the present study, we examine the relationship between students' perceptions of their actual course instructors' presentation fluency (i.e., "instructor fluency") and the students' ratings of their learning for the course material (as well as measures such as their interest and motivation in the course and with their actual grades earned). This relationship has been examined in the laboratory (e.g., Carpenter, Willford, Kornell, & Mullaney, 2013), but has yet to be examined in actual courses. Critically, students' use of instructor fluency as a basis for judging their learning in a course may bias their judgments and impair their study behaviors, for example by producing an illusion of learning (Serra & Metcalfe, 2009).

1.1. Metacognitive Monitoring

Metacognitive judgments are judgments about the status of another cognitive process that people use to control the continued performance of that process (for a brief review, see Serra & Metcalfe, 2009). For example, students engage in metacognitive monitoring when evaluating how well they know or understand lesson content, and they can make explicit metacognitive judgments about the status of their learning (e.g., a judgment that one understands 60% of the content for an upcoming exam; a prediction that one will earn a grade of B+ on an upcoming exam). Related, students can use such judgments as the basis for acts of metacognitive control (e.g., deciding for how long to continue to study; deciding whether or not to stop studying altogether).

When people make such judgments, however, they do not have direct access to the current state of their cognitions to inform those judgments. Rather, they must form their judgments by making inferences based on metacognitive cues. These cues stem from information and experiences they believe are related to that cognition (i.e., Dunlosky & Tauber, 2014; Koriat, 1997; Serra & Metcalfe, 2009). For example, participants in a laboratory-based memory experiment might be asked to make metacognitive judgments about the strength of their memory for each item they studied (i.e., to judge the likelihood that they would recall each one on a subsequent memory test). Participants use a variety of implicit/experiential and explicit/analytical cues such as item relatedness or difficulty (e.g., Koriat & Bjork, 2005), item familiarity (e.g., Reder & Ritter, 1992), and past-test performance (e.g., Serra & Ariel, 2014) when judging their memory for such items. Flaws in students' metacognitive monitoring such as those created by the use of faulty heuristics can impair their study behaviors (Serra & Metcalfe, 2009). Therefore, in the present study, we were concerned with student's use of perceived instructor fluency as a cue for judging their own learning in an actual course.

1.2. Fluency and Learning

The general experience of cognitive ease or processing fluency is associated with the occurrence of a variety of cognitive outcomes such as fast, low-effort, or high-accuracy cognitive processing (for reviews, see Jacoby, Kelley, & Dywan, 1989; Reber, Schwarz, & Winkielman, 2004). Concordantly, the experience of high fluency (i.e., ease) is often associated with a positive interpretation of the state of underlying cognitive processes (e.g., Fernandez-Duque, Baird, & Posner, 2000; Ramachandran & Hirstein, 1999; Vallacher & Nowak, 1999; Winkielman, Schwarz, Fazendeiro, & Reber, 2003); when people experience a cognitive task as being fluent, they often use that experience as a cue to evaluate their performance and infer that they are doing well at it.

The tendency to associate fluency with accurate performance also occurs for students' self-evaluations of their learning. Although the experience of fluency in learning situations correlates with a sense (or the explicit

judgment) that information has been easily acquired, it is not always correlated with actual learning (e.g., Bertsch, Pesta, Wiscott, & McDaniel, 2007; Bjork, 1994; Carpenter et al., 2013; Carpenter, Mickes, Rahman, & Fernandez, 2016; Eitel, Kuhl, Scheiter, & Gerjets, 2014; Hirshman & Bjork, 1988; Leutner, Leopold, & Sumfleth, 2009; Rhodes & Castel, 2008; Sanchez & Khan, 2016; Slamecka & Graf, 1978; Yue, Castel, & Bjork, 2013). For this reason, if students base metacognitive evaluations of their learning on how fluently or easily they processed learning materials—but that experience of fluency is not correlated with their actual learning—then their self-assessments of their learning will be inaccurate. For example, much research has examined the effect of perceptual fluency on participants' memory judgments and demonstrated that physically increasing the fluency of item processing (e.g., making the font of text materials larger or easier to read) affects memory judgments without affecting memory performance: participants tend to judge the more-fluent items (i.e., large-font items) as being more memorable than the less-fluent items (i.e., small-font items), even though memory is not affected by such font-size manipulations (e.g., Magreehan, Serra, Schwartz, & Narciss, 2016; Mueller, Dunlosky, Tauber, & Rhodes, 2014; Rhodes & Castel, 2008).

To date, most experiments examining the effects of fluency on metacognitive judgments of learning have utilized highly-contrived and artificial formatting scenarios such as presenting study materials in a clear versus blurred font (Yue et al., 2013), in an upright versus inverted font (Sungkhasettee, Friedman, & Castel, 2011), and in a high versus low figure-ground contrast (Werth & Strack, 2003). Further, such published effects are commonly obtained using within-participants manipulations of fluency but do not occur when the same manipulations are used between-participants (e.g., Magreehan et al., 2016; Yue et al., 2013). These findings suggest that the effects might reflect a demand characteristic, the application of explicit beliefs rather than a true difference in the experience of fluency or disfluency (cf. Mueller et al., 2014), or extreme situations that would never occur for students (i.e., Magreehan et al., 2016; Serra, 2016a, 2016b). Nevertheless, researchers have argued that such effects pose a threat for learning situations outside of the laboratory (e.g., Finn & Tauber, 2015).

Given these concerns, we did not examine the effects of perceptual fluency on judgments of learning in such ways in the present study. Instead, we examined the relationship between students' perceptions of the fluency of their instructors (i.e., instructor fluency) and their ratings of their learning in an actual course (cf. Williams & Ceci, 1997). For the present purposes, we define *instructor fluency* as the sense of ease that the learner experiences when viewing a lesson that stems from content-independent attributes of the instructor and their presentation of the information. We admit that instructor fluency might involve some aspects of fluency that are akin to perceptual fluency (e.g., how loudly the instructor speaks when teaching; formatting choices in a PowerPoint presentation), but we view the experience of instructor fluency as a larger and more diverse experience when compared to something very specific such as differences in the font formatting of text-based learning materials. We also assume that students will attribute any experience of fluency or disfluency in this regard to the instructor rather than to learning materials or to the course itself.

Carpenter et al. (2013) recently conducted a series of experiments to examine the effects of instructor fluency on participants' ratings of their learning and the quality of the instructor. They designed a fluent and disfluent version of a 65-second video-recorded lesson on the gender of calico cats. Both versions used the same instructor and the same script, but differed in how fluently the instructor presented the information. In the fluent version, the instructor demonstrated good physical posture, used relevant hand gestures, made good eye contact (with the camera), and spoke confidently without referring to notes. In the disfluent version, the instructor demonstrated poor physical posture (slumping), did not use relevant hand gestures, did not make good eye contact, and spoke haltingly while referring to notes. Participants viewed one of the two versions of the video, judged their learning from the video, rated various aspects of the instructor, and then completed a test over the material. Instructor fluency did not affect participants' learning (test performance) from the videos, but it did affect their judgments: participants who viewed the fluent version judged their learning to be higher and judged the instructor to be better than did those who viewed the disfluent version, even though the lesson content and actual learning did not differ by fluency. Similarly, Sanchez & Khan (2016) examined participants' learning and judgments of learning for an online lesson (five minutes in length) narrated by an instructor who was either a Native English speaker or a non-native English speaker with a Mandarin Chinese accent. Much as in Carpenter et al.'s (2013) experiments, although participants scored the same on a post-test regardless of which instructor narrated the lesson (indicating equivalent learning), participants in those experiments judged that they learned more from the Native English instructor than from the instructor with the accent, and also judged that the Native English instructor was a better teacher than was the instructor with the accent. Together, these two sets of experiments

demonstrate that participants' judgments of their learning and of the quality of their instructor can easily be biased by aspects of the instructor that affect the participants' experience of fluency.

The question that motivated the present study, however, is whether similar effects would occur for students in actual courses. [Carpenter et al. \(2016\)](#) recently replicated the results of [Carpenter et al. \(2013\)](#) with longer videos (22 minutes in length), but a single longer video does not even convey the same information as one full lecture in an actual course, let alone the same information conveyed in an entire semester. Further, students in actual classroom learning situations likely have a variety of other additional factors besides instructor fluency on which to base self-assessments of their own learning (e.g., their interactions with the instructor outside of the classroom such as at office hours or via email, their performance on exams and assignments throughout the semester, and even experiencing the same instructor demonstrating a natural range of instructor fluency across days and topics). As such, they might not show any relation between their general experience of instructor fluency and their judgments of learning (but see [Williams & Ceci, 1997](#), who found that differences in instructor expressiveness across semesters affected students' judgments of their learning and evaluations of their instructor in an actual course). Another possibility, however, is that the relationship between instructor fluency and students' judgments about their own learning or the quality of their course or instructor might be so strong that such a relationship occurs even in actual, semester-long courses.

1.3. The Present Study

To examine whether instructor fluency is related to students' evaluations of their learning in an actual course, at the end of a semester we had students rate the fluency of their instructors and judge their own learning for the course materials (we also had them rate several other aspects of their instructor and of the course itself per [Carpenter et al., 2013](#)). As such, we were able to perform a conceptual replication of [Carpenter et al.'s \(2013\)](#) experiments using actual students and instructors. We did not attempt to manipulate fluency across the instructors because, as suggested by [Williams & Ceci \(1997\)](#), it seems unethical to purposely provide a better learning situation to some students than to others for the purposes of experimental research. Instead, we relied on the assumptions that 1) different instructors would naturally differ in instructor fluency; 2) different students would perceive different levels of fluency even for the same instructor; and 3) students' perceptions of instructor fluency matter more for the present purposes than does an instructor's "actual" fluency.

Based on past findings from more contrived laboratory-based experiments (e.g., [Rhodes & Castel, 2008](#); [Sungkhassetee et al., 2011](#); [Werth & Strack, 2003](#); [Yue et al., 2013](#)) and cue-utilization theories of metacognitive monitoring ([Dunlosky & Tauber, 2014](#); [Koriat, 1997](#)), one prediction is that students *will* use their experience of instructor fluency during the semester as a major basis for their judgments of learning and other judgments about their instructor at the end of the semester. Accordingly, we would predict that students' sense of instructor fluency would correlate with their judgments of learning and other ratings of the instructor and the course (cf. [Carpenter et al., 2013](#); [Williams & Ceci, 1997](#)). In contrast, another prediction is that students will not use their experience of instructor fluency during the semester as a major basis for any of their judgments or ratings. By the end of the semester, students should have a wealth of other information besides their experience of instructor fluency to use as a cue for their judgments and ratings. [Magreehan et al. \(2016\)](#) demonstrated that even for very simple memory materials such as word pairs, the presence of relevant memory cues such as item relatedness overshadowed perceptual-fluency effects on participants' learning judgments; an actual semester-long course likely contains an even greater wealth of informative cues about learning that students might instead consult to judge their learning. Accordingly, we would predict that students' sense of instructor fluency would not correlate with their judgments of learning and other ratings of the instructor and the course.

2. Method

2.1. Participants

The participants were 606 undergraduate students enrolled in an introductory psychology course at Texas Tech University. All students were enrolled during the same semester, but were spread across 25 different sections. The sample of student respondents was 66% female and 34% male with a modal age of 18 years old ($M = 18.9$ years old, $SD = 1.2$). Including those who belonged to more than one race or ethnic group, the sample was 73% White or European American, 20% Hispanic or Latino, 7% Black or African American, 4% Asian or Asian American, 1% Native American or Alaskan Native, and 2% who indicated "other". The sample was 70% fresh-

men, 19% sophomores, 6% juniors, and 5% seniors. All students received course credit for their participation.

The 25 course sections were taught by 14 different instructors (11 instructors taught two sections each; 3 instructors taught one section each). Nevertheless, there was high consistency across these instructors' approach to the course, as many of the course's attributes were pre-determined by departmental policy. All of the instructors were graduate students in psychology who had completed at least 18 credit hours of graduate coursework in psychology. All instructors were required to use the same textbook and had access to the same associated resources (i.e., the publisher's PowerPoint files, multimedia resources, test bank, etc.). Most instructors structured their course similarly: no large paper assignments, a research-participation requirement, and three or four non-cumulative exams throughout the semester that consisted primarily of multiple-choice questions.

2.2. Materials

The materials for the present study were an online survey that we administered to students enrolled in an introductory psychology course at the end of the semester. This survey included the following subsets of questions.

2.2.1. Instructor Fluency

To assess instructor fluency, we created nine Likert-scale questions for participants to answer about their instructor (Table 1). We based these questions on Carpenter et al.'s (2013) descriptions of their fluent and disfluent instructor conditions and on other attributes that might distinguish fluent from disfluent instructors. These questions asked about the instructor's ability to explain information in a meaningful way, to speak clearly, to hold the students' attention, to maintain eye contact with students, to respond to student questions, to maintain good body posture, to use PowerPoint and other visual aids, to use handouts and other supplements, and to speak loudly. All questions had five responses. We discuss the reliability of this measure in the Results section.

Table 1. Questions (with response options) to assess instructor fluency.

Question				
In general, how well did your instructor explain information so the class could understand it?				
Not well at all	Somewhat well	Moderately well	Very well	Extremely well
In general, how clearly did your instructor speak when teaching or addressing the class?				
Not clearly at all	Somewhat clearly	Moderately clearly	Very clearly	Extremely clearly
In general, how well did your instructor maintain the students' attention when teaching or addressing the class?				
Not well at all	Somewhat well	Moderately well	Very well	Extremely well
In general, how well did your instructor maintain eye contact with students when teaching or addressing the class?				
Not well at all	Somewhat well	Moderately well	Very well	Extremely well
In general, how well did your instructor answer students' questions when teaching or reviewing?				
Not well at all	Somewhat well	Moderately well	Very well	Extremely well
In general, how well did your instructor maintain good body posture when teaching or addressing the class?				
Not well at all	Somewhat well	Moderately well	Very well	Extremely well
In general, how well did your instructor make use of visual aids such as PowerPoint or writing on the board when teaching?				
Not well at all	Somewhat well	Moderately well	Very well	Extremely well
In general, how well did your instructor make use of provided materials such as handouts or outlines when teaching?				
Not well at all	Somewhat well	Moderately well	Very well	Extremely well
In general, how loudly did your instructor speak when teaching or addressing the class?				
Not loudly at all	Somewhat loudly	Moderately loudly	Very loudly	Extremely loudly

Note. Participants answered the nine questions in a random order. Responses were scored from 1 to 5, with 5 indicating greatest fluency. Response scorings were totaled to yield a total fluency score.

2.2.2. Objective Learning

Participants granted us permission to access their actual final numerical grade in the course (i.e., their final grade ranging from 0% to 100%) as an objective measure of their learning in the course.

2.2.3. Subjective Learning

We created two questions to assess subjective learning outcomes in the course and to examine whether students' learning judgments were related to their experience of instructor fluency (cf. Carpenter et al., 2013). Participants indicated what percentage of the course content they felt they had learned (on a scale from 0% to 100%) and what numerical grade they expected to earn in the course (on a scale from 0% to 100%).

2.2.4. Instructor Efficacy

We adapted four questions from Carpenter et al. (2013) in order to compare our results to theirs and to examine whether students' perceptions of the quality of their instructors was related to their experience of instructor fluency. These four questions asked participants to rate their instructor's organization, preparedness, knowledge, and effectiveness via five-point Likert scale questions (e.g., "Not at all effective", "Somewhat effective", "Moderately effective", "Very effective", "Extremely effective"). Given that one goal of the present study was to replicate Carpenter et al.'s lab-based results with actual students in an actual course, we did not greatly alter the nature of these questions or use multiple questions to assess what Carpenter et al. assessed with one question.

2.2.5. Topic Interest

We adapted two questions from Carpenter et al. (2013) in order to compare our results to theirs and to examine whether students' ratings of their interest in the topic of the course was related to their experience of instructor fluency. These two questions asked participants to rate how interested they were in the topics covered in the course and how motivated they were to learn those topics (cf. Carpenter et al., 2013), also via five-point Likert scale questions (e.g., "Not at all motivated", "Somewhat motivated", "Moderately motivated", "Very motivated", "Extremely motivated"). Again, we purposely did not greatly alter the nature of these questions or use multiple questions to assess what Carpenter et al. assessed with one question.

2.2.6. Course Evaluation

We included three Likert-scale questions to assess students' subjective assessment of the quality of their instructor and the course and to examine whether students' perceptions of these attributes were related to their experience of instructor fluency. These three questions asked students to rate their instructor, the course, and the course with their instructor, each on a five-point scale (e.g., "Terrible", "Poor", "Average", "Good", "Excellent"). We designed these questions to mimic actual course evaluation questions used at many universities.

2.3. Procedure

Participants completed the present measures as part of an online survey that we administered at the end of a full-length semester. Students either gave informed consent at the outset of the survey or opted out of completing the survey. Students also had the option to stop completing the survey at any time, and to skip any questions they did not want to answer, without penalty. After participants gave consent to participate, they answered the present questions in a randomized order within blocks of questions presented in a fixed order. Other questions in the survey were not relevant for the present purposes, so we do not report results from them here. Students also provided us with permission to access their final course grades (or denied us permission to do so). We obtained their final course grades from the instructors, and then de-identified the data for analysis.

2.4. Statistical Analysis

Descriptive analyses for the present data set involved calculating means and standard deviations across participants for each of the present scale measures. To consider the internal reliability of the instructor fluency questionnaire, we conducted a principal-component factor analysis of those questions. To consider the test-retest reliability of this questionnaire, we compared instructor fluency scores for eight instructors across the present and subsequent semesters using a Spearman rank-order correlation. Most important, the focal analyses of this study involved calculating Pearson r correlations between the present scale measures. We did this using both the raw

data and by calculating partial correlations after controlling for students' actual grades in the course.

3. Results

3.1. Instructor Fluency

First, we converted participants' responses to the instructor-fluency questions into numerical scores ranging from 1 (i.e., "Not very well") to 5 (i.e., "Extremely well") and calculated Pearson r correlations across participants' responses. Their responses to the nine questions were all correlated with each other at the $p < 0.001$ level (Table 2). As such, it seems that these questions tapped a similar construct even though each question asked about a different aspect of the instructors' teaching. We then summed participants' responses to the nine questions to yield a total instructor-fluency score for each participant. The values for this outcome could range from 9 to 45, and the range for this measure in the actual sample was 10 to 45. Not surprisingly, participants' total fluency scores were correlated with their responses to all nine of the questions in the instructor-fluency questionnaire (Table 2). These Pearson r correlations were all significant at the $p < 0.001$ level.

3.1.1. Questionnaire Factor Analysis

We conducted a principal-component factor analysis of the nine questions that we created to assess instructor fluency. As we report in Table 2, responses to the nine questions were all significantly correlated. The KMO index was 0.94 (indicating "superb" sampling adequacy) and Bartlett's test was significant at the $p < 0.001$ level (rejecting the null hypothesis of an identity matrix). The communality extractions (Table 3) were all 0.3 or greater except for that of how loudly the instructor spoke, which was just under 0.3 (confirming that the items shared common variance with the other items).

Our factor analysis strongly supported the idea that summing participants' responses to our questions together to yield a "total instructor fluency" score was appropriate. All nine questions loaded heavily on a single factor. We report these loadings (from highest to lowest) in Table 3. This factor had an Eigen value of 5.3 and explained 59.4% of the variance. Eigen values immediately fell below 1.0 with additional factors, so we maintained the single-factor interpretation. No rotation was possible.

3.1.2. Questionnaire Reliability

In order to assess the reliability of this measure, we estimated the internal consistency of our instructor-fluency scale using Cronbach's α . The α value for the nine questions was 0.91, indicating very high internal consistency (at least with the present sample).

Further, using the same instructor-fluency questionnaire, we collected data from new students enrolled in the same course during the following semester. Eight instructors who taught the course during the present semester

Table 2. Correlations between total fluency and subcomponent ratings.

Measure	Descriptives		Correlations										
	M	SD	1	2	3	4	5	6	7	8	9	10	
1. Fluency Total	37.0	6.2	–										
2. Explain Information	4.2	0.9	0.87	–									
3. Speak Clearly	4.2	0.9	0.85	0.75	–								
4. Maintain Attention	4.0	1.0	0.84	0.73	0.73	–							
5. Make Eye Contact	4.3	0.8	0.81	0.67	0.66	0.65	–						
6. Answer Questions	4.3	0.9	0.80	0.72	0.64	0.63	0.63	–					
7. Good Body Posture	4.2	0.9	0.77	0.61	0.60	0.58	0.65	0.58	–				
8. Use of Visual Aids	4.5	0.7	0.70	0.56	0.56	0.50	0.55	0.60	0.52	–			
9. Use of Handouts	3.7	1.2	0.68	0.51	0.48	0.48	0.46	0.44	0.45	0.38	–		
10. Speak Loudly	3.5	0.8	0.56	0.40	0.45	0.41	0.38	0.33	0.31	0.28	0.30	–	

Note. All correlations were significant at the $p < 0.001$ level. See Table 1 for the full wording of each question.

Table 3. Questionnaire extractions and factor loadings.

Question	Extraction	Factor Loading
Explain Information	0.77	0.88
Speak Clearly	0.74	0.86
Maintain Attention	0.70	0.84
Make Eye Contact	0.68	0.83
Answer Questions	0.67	0.82
Good Body Posture	0.60	0.77
Use of Visual Aids	0.52	0.72
Use of Handouts	0.41	0.64
Speak Loudly	0.28	0.53

Note. Participants' responses to all but one question of the instructor-fluency questionnaire had an extraction above 0.3 ("Extraction" column). Participants' responses to the nine questions of the instructor-fluency questionnaire all loaded heavily on a single factor ("Factor Loading" column).

taught it again in the following semester, so we were able to compare their mean total fluency rating across the two semesters. As can be seen in **Table 4**, instructors received largely the same mean rating across the semesters despite the fact that they were teaching the course anew and to a new set of students in the second semester. The rank-ordering of the instructors by their fluency rating in the two semesters was highly consistent: we calculated a Spearman rank-order correlation coefficient across the two semesters' ratings (**Table 4**), and they were highly correlated, $\rho = 0.81$, $p = 0.02$. For example, the same instructor was rated as most-fluent in both semesters and the same instructor was rated as least-fluent in both semesters. In short, our instructor-fluency scale has good internal reliability and good test-retest reliability.

3.2. Objective Learning

Participants' total fluency scores were correlated with their actual final grades in the course (**Table 5**). Total fluency was correlated with participants' actual grades, but this correlation was small ($r = 0.13$, $p < 0.001$).

3.3. Subjective Learning

Participants' total fluency scores were correlated with their estimates of what percentage of the course material they had learned and with their final-grade predictions questions (center of **Table 5**). Given the correlation of total fluency with actual final grade in the course, we felt it was important to demonstrate that the present ratings have psychological independence from the actual grades students earned in the course. To examine this, we calculated the present correlations after controlling for students' actual grades in the course (right column of **Table 5**). Except for the correlation between total fluency and predicted grade, the relationship between total fluency and all other judgments maintained after controlling for actual grades in the course (right column of **Table 5**).

3.4. Instructor Efficacy

We converted participants' responses to each instructor-efficacy question into a numerical score ranging from 1 (e.g., "Not very organized") to 5 (i.e., "Extremely organized") and calculated Pearson r correlations across participants' responses. Total fluency was correlated with participants' ratings of their instructor's organization, preparedness, knowledge, and effectiveness (center of **Table 5**). These relationships maintained after controlling for actual grades in the course (right column of **Table 5**).

3.5. Topic Interest

We converted participants' responses to each question into a numerical score ranging from 1 (e.g., "Not at all motivated") to 5 (i.e., "Extremely motivated") and calculated Pearson r correlations across participants' responses. Total fluency was correlated with participants' ratings of how interested they were in the topics covered

Table 4. Mean total fluency ratings for eight instructors in the present and following semesters.

Identifier	Present Semester			Following Semester		
	n	M	SD	n	M	SD
Instructor 1	63	40.9	3.6	65	40.5	3.5
Instructor 2	45	39.7	3.8	71	38.1	5.4
Instructor 3	18	39.6	4.1	30	38.7	4.3
Instructor 4	63	38.4	5.1	55	38.5	5.0
Instructor 5	69	37.5	4.8	78	38.6	4.8
Instructor 6	29	35.8	5.0	49	36.6	6.7
Instructor 7	39	35.2	5.8	14	36.8	6.7
Instructor 8	30	33.4	6.9	43	34.4	5.5

Note. Values are the mean fluency rating for eight instructors who taught the same course in the present semester (during which we collected the present data) and in the following semester. We did not analyze the data from the following semester except for the purpose of this comparison.

Table 5. Correlations between total fluency and other ratings.

Measure	Descriptive Statistics		Correlations to Fluency		Partial Correlations	
	M	SD	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Fluency Total	37.0	6.2	–	–	–	–
Objective Learning Rating						
Actual Grade	89.8	10.2	0.13	0.001	–	–
Subjective Learning Ratings						
Learning Judgment	80.8	15.9	0.40	<0.001	0.38	<0.001
Predicted Grade	88.6	7.8	0.12	0.003	0.0	0.4
Instructor-Efficacy Ratings						
Organized	4.2	0.8	0.70	<0.001	0.70	<0.001
Prepared	4.3	0.8	0.72	<0.001	0.72	<0.001
Knowledge	4.3	0.8	0.71	<0.001	0.71	<0.001
Effective	4.1	0.9	0.81	<0.001	0.81	<0.001
Topic-Interest Ratings						
Interest	3.8	1.0	0.36	<0.001	0.34	<0.001
Motivation	3.7	1.0	0.45	<0.001	0.44	<0.001
Course-Evaluation Ratings						
Instructor Rating	4.4	0.8	0.75	<0.001	0.74	<0.001
Course Rating	4.0	0.8	0.40	<0.001	0.38	<0.001
Course with Instructor	4.4	0.8	0.72	<0.001	0.72	<0.001

Note. Values are mean rating for each measure (left column), Pearson correlations between total fluency ratings and the other present measures (center column), and these same correlations after controlling for students’ actual grades in the course (right column).

in the course and how motivated they were to learn those topics (center of **Table 5**). These relationships maintained after controlling for actual grades in the course (right column of **Table 5**).

3.6. Course Evaluation

We converted participants’ responses to each question into a numerical score ranging from 1 (i.e., “Terrible”) to

5 (i.e., “Excellent”). Total fluency was correlated with participants’ ratings of their instructor, of the course, and of the course with that instructor (center of **Table 5**). These relationships maintained after controlling for actual grades in the course (right column of **Table 5**). Importantly, the correlation between total fluency and students’ course rating was of smaller magnitude than were the correlations between total fluency and students’ ratings of both the instructor and the course with that instructor. This suggests that students’ total fluency scores reflect their experience of the instructor more so than their satisfaction with the course or topic.

4. Discussion

Despite a growing body of research demonstrating that perceptual-fluency manipulations can affect students’ judgments of their learning (e.g., [Mueller et al., 2014](#); [Sungkhasettee et al., 2011](#); [Werth & Strack, 2003](#); [Yue et al., 2013](#)), few if any research studies have examined such effects in more realistic settings or conditions (i.e., either in the classroom or using realistic variations in perceptual fluency; see [Magreehan et al., 2016](#); [Serra, 2016a, 2016b](#) for further criticisms). Research examining the effects of instructor fluency rather than perceptual fluency has utilized more realistic variations in students’ experience of fluency, but the setting has remained artificial (i.e., very short instructional videos with no extrinsic investment on the part of the participants; e.g., [Carpenter et al., 2013](#); [Carpenter et al., 2016](#); [Sanchez & Khan, 2016](#); but see [Williams & Ceci, 1997](#)). To this end, the purpose of the present study was to determine whether instructor fluency is related to students’ evaluations of their learning in an actual course, much as occurs in more contrived laboratory experiments that have examined this question (e.g., [Carpenter et al., 2013](#); [Carpenter et al., 2016](#); [Sanchez & Khan, 2016](#)).

We found that students’ ratings of instructor fluency were correlated with their judgments of learning, various ratings of their instructor, and various ratings of the course and its topics. Importantly, these relationships maintained when we controlled for student’s actual performance in the course and occurred even though students had a semester’s worth of experience with the course and instructor (i.e., assignments; grades; within-instructor variations in fluency) that they could consider when making their judgments. These results suggest that instructor fluency is a major source of information that students factor into such judgments in both the laboratory and in the classroom. The present study therefore provides an important link between highly-contrived laboratory examinations of fluency effects and more genuine examinations of the relationship between fluency and students’ judgments of their learning, their instructors, and their courses.

The present findings are perhaps even more surprising when we consider that, in most concordant laboratory studies, participants made their judgments immediately after being exposed to either fluent or disfluent presenters or study materials. When the making of the judgment is delayed from the exposure to the materials in the laboratory, however, judgments of learning do not seem to show fluency effects (e.g., [Hu, Liu, Li, & Luo, 2016](#)). In contrast, participants in the present study made their judgments outside of the classroom and at a long temporal delay from most of their experience with the fluency of their instructors. Nevertheless, participants’ judgments in the present study were related to instructor fluency despite how much time had elapsed between their exposure to their instructors and their making of the present ratings. This suggests that instructor fluency in the classroom might exert a stronger or longer-lasting influence on students’ judgments of learning than might other forms of fluency (i.e., perceptual fluency).

Implications and Future Directions

Problematically, the relationship between instructor fluency and students’ judgments of their learning in actual courses seems very strong; the relationship occurs even though students likely have numerous other sources of information they can consult to judge their learning (or to rate the quality of their instructors and courses) and maintain when we control for actual grades earned in the course. If further research demonstrates that the experience of instructor fluency can impair the efficacy of students’ study behaviors in their courses, then applied researchers may have difficulty identifying methods to reduce students’ use of this heuristic. As we previously note, the general heuristic that the experience of fluency is associated with positive performance is pervasive (e.g., [Fernandez-Duque et al., 2000](#); [Jacoby et al., 1989](#); [Ramachandran & Hirstein, 1999](#); [Reber et al., 2004](#); [Vallacher & Nowak, 1999](#); [Winkielman et al., 2003](#)), so applied researchers may have to work particularly hard to eliminate its use by students in the context of learning.

In actual courses, the experience of high instructor fluency can lead students to overestimate their level of learning and under-prepare for exams (but see [Carpenter et al., 2013](#), Experiment 2), make poor restudy deci-

sions (cf. Shanks & Serra, 2014), or even change their academic major (Stinebrickner & Stinebrickner, 2014). Related, research indicates that students' likelihood of communicating with instructors outside of the classroom (e.g., emailing an instructor with a question; attending office hours) is negatively correlated with their perception of instructor clarity (Sidelinger, Bolen, McMullen, & Nyeste, 2015). Given such findings, researchers should identify relationships between instructor fluency and students' study behaviors. Researchers should also consider whether instructor fluency biases students' actual evaluations of their instructors and courses.

References

- Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The Generation Effect: A Meta-Analytic Review. *Memory & Cognition*, 35, 201-210. <http://dx.doi.org/10.3758/BF03193441>
- Bjork, R. A. (1994). Memory and Metamemory Considerations in the Training of Human Beings. In J. Metcalfe, & A. Shimamura (Eds.), *Metacognition: Knowing about Knowing* (pp. 185-205). Cambridge, MA: MIT Press.
- Carpenter, S. K., Mickes, L., Rahman, S., & Fernandez, C. (2016). The Effect of Instructor Fluency on Students' Perceptions of Instructors, Confidence in Learning, and Actual Learning. *Journal of Experimental Psychology: Applied*, in press. <http://dx.doi.org/10.1037/xap0000077>
- Carpenter, S. K., Wilford, M. M., Kornell, N., & Mullaney, K. M. (2013). Appearances Can Be Deceiving: Instructor Fluency Increases Perceptions of Learning without Increasing Actual Learning. *Psychonomic Bulletin & Review*, 20, 1350-1356. <http://dx.doi.org/10.3758/s13423-013-0442-z>
- Duarte, N. (2012). *HBR Guide to Persuasive Presentations* (pp. 155-160). Harvard Business Press.
- Dunlosky, J., & Tauber, S. K. (2014). Understanding People's Metacognitive Judgments: An Ismechanism Framework and Its Implications for Applied and Theoretical Research. In T. Perfect, & S. Lindsay (Eds.). *Handbook of Applied Memory* (pp. 444-464). Thousand Oaks, CA: Sage. <http://dx.doi.org/10.4135/9781446294703.n25>
- Eitel, A., Kuhl, T., Scheiter, K., & Gerjets, P. (2014). Disfluency Meets Cognitive Load in Multimedia Learning: DOES harder-to-Read Mean Better-to-Understand? *Applied Cognitive Psychology*, 28, 488-501. <http://dx.doi.org/10.1002/acp.3004>
- Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive Attention and Metacognitive Regulation. *Consciousness and Cognition*, 9, 288-307. <http://dx.doi.org/10.1006/ccog.2000.0447>
- Finn, B., & Tauber, S. K. (2015). When Confidence Is Not a Signal of Knowing: How Students' Experiences and Beliefs about Processing Fluency Can Lead to Miscalibrated Confidence. *Educational Psychology Review*, 27, 567-586. <http://dx.doi.org/10.1007/s10648-015-9313-7>
- Hirshman, E., & Bjork, R. A. (1988). The Generation Effect: Support for a Two-Factor Theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 14, 484-494. <http://dx.doi.org/10.1037/0278-7393.14.3.484>
- Hu, X., Liu, Z., Li, T., & Luo, L. (2016). Influence of Cue Word Perceptual Information on Metamemory Accuracy in Judgment of Learning. *Memory*, in press
- Jackson-Beeck, M., & Meadow, R. G. (1979). Content Analysis of Televised Communication Events: The Presidential Debates. *Communication Research*, 6, 321-344. <http://dx.doi.org/10.1177/009365027900600304>
- Jacoby, L., Kelley, C., & Dywan, J. (1989). *Memory Attributions. Varieties of Memory and Consciousness: Essays in Honour of Endel Tulving* (pp. 391-422). Hillsdale, NJ: Erlbaum.
- Koriat, A. (1997). Monitoring One's Own Knowledge during Study: A Cue-Utilization Approach to Judgments of Learning. *Journal of Experimental Psychology: General*, 126, 349-370. <http://dx.doi.org/10.1037/0096-3445.126.4.349>
- Koriat, A., & Bjork, R. A. (2005). Illusions of Competence in Monitoring One's Knowledge during Study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 187-194. <http://dx.doi.org/10.1037/0278-7393.31.2.187>
- Leutner, D., Leopold, C., & Sumfleth, E. (2009). Cognitive Load and Science Text Comprehension: Effects of Drawing and Mentally Imaging Text Content. *Computers in Human Behavior*, 25, 284-289. <http://dx.doi.org/10.1016/j.chb.2008.12.010>
- Magreehan, D. A., Serra, M. J., Schwartz, N. H., & Narciss, S. (2016). Further Boundary Conditions for the Effects of Perceptual Disfluency on Judgments of Learning. *Metacognition and Learning*, 11, 35-56. <http://dx.doi.org/10.1007/s11409-015-9147-1>
- Mueller, M. L., Dunlosky, J., Tauber, S. K., & Rhodes, M. G. (2014). The Font-Size Effect on Judgments of Learning: Does It Exemplify Fluency Effects or Reflect People's Beliefs about Memory? *Journal of Memory and Language*, 70, 1-12. <http://dx.doi.org/10.1016/j.jml.2013.09.007>
- Naftulin, D. H., Ware Jr., J. E., & Donnelly, F. A. (1973). The Doctor Fox Lecture: A Paradigm of Educational Seduction. *Academic Medicine*, 48, 630-635. <http://dx.doi.org/10.1097/00001888-197307000-00003>
- Ramachandran, V. S., & Hirstein, W. (1999). The Science of Art: A Neurological Theory of Aesthetic Experience. *Journal*

- of *Consciousness Studies*, 6, 15-51.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience? *Personality and Social Psychology Review*, 8, 364-382. http://dx.doi.org/10.1207/s15327957pspr0804_3
- Reder, L. M., & Ritter, F. E. (1992). What Determines Initial Feeling of Knowing? Familiarity with Question Terms, Not with the Answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 435-451. <http://dx.doi.org/10.1037/0278-7393.18.3.435>
- Rhodes, M. G., & Castel, A. D. (2008). Memory Predictions Are Influenced by Perceptual Information: Evidence for Meta-Cognitive Illusions. *Journal of Experiment Psychology: General*, 137, 615-625. <http://dx.doi.org/10.1037/a0013684>
- Sanchez, C. A., & Khan, S. (2016). Instructor Accents in Online Education and Their Effect on Learning and Attitudes. *Journal of Computer Assisted Learning*. <http://dx.doi.org/10.1111/jcal.12149>
- Serra, M. J. (2016a). Does Processing Fluency Really Matter for Metacognition in Actual Learning Situations? Part One: Fluency in the Laboratory. <http://www.improvewithmetacognition.com/>
- Serra, M. J. (2016b). Does Processing Fluency Really Matter for Metacognition in Actual Learning Situations? Part Two: Fluency in the Classroom. <http://www.improvewithmetacognition.com/>
- Serra, M. J., & Ariel, R. (2014). People Use the Memory for Past-Test Heuristic as an Explicit Cue for Judgments of Learning. *Memory & Cognition*, 42, 1260-1272. <http://dx.doi.org/10.3758/s13421-014-0431-0>
- Serra, M. J., & Metcalfe, J. (2009). Effective Implementation of Metacognition. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition and Education* (pp. 278-298). New York: Routledge.
- Shanks, L. L., & Serra, M. J. (2014). Domain Familiarity as a Cue for Judgments of Learning. *Psychonomic Bulletin & Review*, 21, 445-453. <http://dx.doi.org/10.3758/s13423-013-0513-1>
- Sidelinger, R. J., Bolen, D. M., McMullen, A. L., & Nyeste, M. C. (2015). Academic and Social Integration in the Basic Communication Course: Predictors of Students' Out-of-Class Communication and Academic Learning. *Communication Studies*, 66, 63-84. <http://dx.doi.org/10.1080/10510974.2013.856807>
- Slamecka, N. J., & Graf, P. (1978). The Generation Effect: Delineation of a Phenomenon. *Journal of Experimental Psychology: Human Learning & Memory*, 4, 592-604. <http://dx.doi.org/10.1037/0278-7393.4.6.592>
- Stinebrickner, R., & Stinebrickner, T. R. (2014). A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout. *The Review of Economic Studies*, 81, 426-472. <http://dx.doi.org/10.1093/restud/rdt025>
- Sungkhasettee, V. W., Friedman, M. C., & Castel, A. D. (2011). Memory and Metamemory for Inverted Words: Illusions of Competency and Desirable Difficulties. *Psychonomic Bulletin & Review*, 18, 973-978. <http://dx.doi.org/10.3758/s13423-011-0114-9>
- Vallacher, R. R., & Nowak, A. (1999). The Dynamics of Self-Regulation. In R. R. Wyer (Ed.), *Perspectives on Behavioral Self-Regulation: Advances in Social Cognition, Vol. XII* (pp. 241-259). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Werth, L., & Strack, F. (2003). An Inferential Approach to the Knew-It-All-Along Phenomenon. *Memory*, 11, 411-419. <http://dx.doi.org/10.1080/09658210244000586>
- Williams, W. M., & Ceci, S. J. (1997). "How'm I Doing?" Problems with Student Ratings of Instructors and Courses. *Change: The Magazine of Higher Learning*, 29, 12-23. <http://dx.doi.org/10.1080/00091389709602331>
- Winkielman, P., Schwarz, N., Fazendeiro, T. A., & Reber, R. (2003). The Hedonic Marking of Processing Fluency: Implications for Evaluative Judgment. In J. Musch, & K. C. Klauer (Eds.), *The Psychology of Evaluations. Affective Processes in Cognition and Emotion* (pp. 189-217). Mahwah, NJ: Lawrence Erlbaum.
- Yue, C. L., Castel, A. D., & Bjork, R. A. (2013). When Disfluency Is—And Is Not—A Desirable Difficulty: The Influence of Typeface Clarity on Metacognitive Judgments and Memory. *Memory & Cognition*, 41, 229-241. <http://dx.doi.org/10.3758/s13421-012-0255-8>