International Journal of Communication 11(2017), 1683–1704 1932–8036/20170005

Attributional Chromatics: How Does the Color of Written Communication Affect Interpersonal Perceptions?

ADAM S. RICHARDS
Texas Christian University, USA

EDWARD L. FINK
Temple University, USA

This study investigated how the color of written communication influences interpersonal attributions. An experiment was conducted whereby students read a peer's graded essay that varied according to the color of the ink used by the instructor. Feedback in red, compared with green or black, generally elicited more negative perceptions. Students reading feedback in red experienced greater negative emotion, gave the essay a lower grade, and judged the student-author to be less capable. A mediation model whereby negative emotion intervened between the color manipulation and the assessments made toward the essay, the student, and the instructor was supported. Nonrecursivity between student and instructor characteristics was found. Positive perceptions of student ability reduced perceptions of instructor competence, whereas positive perceptions of instructor competence increased perceptions of student ability.

Keywords: attribution, color, emotion, nonverbal communication

Anecdotally, fewer and fewer teachers use red pens because red is thought to upset students (e.g., Turvill, 2014). The purpose of this study is to understand the process by which color used for a written message affects attributions made about the target and source of the communication. We investigate this question in the specific context of the student–instructor relationship by assessing how color, in the form of an instructor's written feedback on students' assignments affects students' perceptions of the graded academic work, their peers, and their instructors.

Colors as nonverbal cues communicate a great deal. For example, studies show that the color of a person's clothing (Elliot & Pazda, 2012) or of a written message (Elliot, Maier, Moller, Friedman, & Meinhardt, 2007) affect behavior. Yet the *process* by which color affects interpersonal behavior has received relatively little attention (Knapp, Hall, & Horgan, 2014). By integrating Elliot and Maier's (2012) color-in-context theory and Weiner's (1986) attributional model of motivation and emotion, we argue that

Adam S. Richards: adam.richards@tcu.edu
Edward L. Fink: elf1@temple.edu
Date submitted: 2016–08–02

the color of written instructional feedback influences interpersonal attributions about students and instructors via initial emotional responses. Our study diverges from previous research (Dukes & Albanesi, 2013) by emphasizing the primary role of emotion within the color–perception relationship. Further, we seek to assess how the attributions made by student-observers who have seen an evaluative communication made by an instructor toward another student are processed. The academic context of instructional feedback allows for a practically important and theoretically relevant study of interpersonal attribution via color cues.

Thus, we frame the present study with regard to two questions. First, what attributions do students make about the quality of an essay, its student-author, and its instructor-grader after receiving feedback made in red? Second, by what process does red feedback affect students' emotional and interpersonal responses? To answer these questions, we conducted an experiment in which students made attributions about a student and instructor after reading a paper with feedback in red, green, or black, or without any feedback.[1]

Theoretical Approach to Color

Color-in-context theory (Elliot & Maier, 2014) explains people's varying responses to color cues. The theory states that colors communicate meanings that influence cognitions and behaviors. Colors can elicit either positive associations, which evoke approach-oriented responses, or negative associations, which evoke avoidance-oriented responses. Colors' effects on psychological functioning frequently go unrecognized because of the brain's automatic and implicit processing of colors (Elliot et al., 2007).

According to Elliot and Maier (2012), a color's meaning is interpreted in the environment in which it is encountered. Contexts that affect color perceptions include additional visual information, such as the shape or texture of the object on which the color is seen. This idea explains why red has different connotations depending on context. In affiliative situations, people wearing red clothing are perceived to be more attractive than those adorned in other colors (e.g., Elliot & Pazda, 2012).

However, in achievement situations, red is associated with negative or avoidant responses (Thorstenson, 2015). Red is associated with failure and green with success in intellectual achievement contexts in which one's competence is evaluated (Moller, Elliot, & Maier, 2009). For example, seeing red on a test booklet resulted in participants choosing easier intellectual tasks (Elliot & Maier, 2012) and physically distancing themselves from the test booklet (Elliot, Maier, Binser, Friedman, & Pekrun, 2009), which indicates avoidance. The color of a test-giver's shirt caused participants to choose easier questions to answer (Tanaka & Tokuno, 2011). Red lowered scores on intelligence tests if it appeared as the color of the test booklet cover (Maier, Elliot, & Lichtenfeld, 2008), particularly for those with little self-control (Bertrams, Baumeister, Englert, & Furley, 2015).The color of the ink used to write an identification number on the cover of the test booklet elicited similar results (Elliot et al., 2007). Red progress bars and forward arrows in Web-based testing diminished cognitive abilities (Gnambs, Appel, & Batinic, 2010). People who received red feedback also performed worse in subsequent trials, making more errors in a

[1] Although technically not a color, black will be labeled a color in this study.

rapid-response task (Houtman & Notebaert, 2013) and making poorer financial decisions (Shavit, Rosenboim, & Cohen, 2013). The effects hold for children: Red screens led to worse scores on reading comprehension tests and on other cognitive tasks (Brooker & Franklin, 2016).

Despite the evidence showing the negative effects of red in achievement contexts, the color–emotion relationship has not been well explained (Elliot & Maier, 2014). We seek to demonstrate that color's influence on cognition is mediated by emotion. Although color-in-context theory acknowledges that color acts as an implicit affective cue, the theory does not designate emotion as necessarily antecedent to cognition. Current theorizing about color has focused on cognitive and behavioral outcomes, but not necessarily on the process by which these outcomes are brought about.

How the color of instructional feedback comes to influence interpersonal perceptions in achievement contexts may be explained by attribution theory, which accounts for the process by which people make causal explanations to observed stimuli (Heider, 1958). Weiner's (1986) attributional model of motivation and emotion, in particular, may account for the process by which color influences interpersonal attributions. The model asserts that people make causal ascriptions about a person's achievement (e.g., ability and effort) and personality subsequent to experiencing unexpected negative emotion from a stimulus. Thus, outcome-dependent affect—affect that is a direct response to a subjective experience—is theorized to precede interpersonal attributions: Only after people experience negative affect do they make causal ascriptions about others. The purpose of the current investigation is to assess whether the effect of instructional feedback color on interpersonal attributions is mediated by negative affect. Because red elicits negative responses in achievement contexts (Elliot & Maier, 2012), the possibility exists for such responses to cause student (under)achievement according the process theorized by Weiner.

In addition to the color red affecting a person's own avoidant responses, research shows that the color red affects evaluative attributions toward other people in achievement contexts for both physical and intellectual tasks. For example, hypothetical sparring opponents who were imagined to wear red were perceived as more dominant and threatening (Feltman & Elliot, 2011). Taekwondo referees assigned greater success to combatants who wore red compared with blue (Hagemann, Strauss, & Leißing, 2008). Boxers wearing red are perceived as more brave and aggressive compared with ones wearing blue (Sorokowski & Szmajke, 2007). Mock negotiators perceived their partners as more intimidating when the others bargained using red tokens as opposed to tokens of other colors (Ten Velden, Baas, Shalvi, Preenen, & De Dreu, 2012). Even red shapes, compared with blue shapes, are considered more dominant, aggressive, and likely to win a physical fight against a black shape (Little & Hill, 2007). Outside of competitive settings, men wearing red are also perceived by others as more aggressive, dominant, and angry (Wiedermann, Burt, Hill, & Barton, 2015).

Red affects the attributions of those who evaluate intellectual performance as well. Rutchick, Slepian, and Ferris (2010) found that graders "using red pens made the concept of errors and poor performance more cognitively accessible" (p. 706), and instructors evaluating excerpts from a student essay using a red pen "marked more errors . . . than did participants using blue pens" (p. 706). Furthermore, graders using red pens (as compared with those using blue pens) gave the excerpt from an

essay that they read a lower grade (Rutchick et al., 2010). If instructors grade more harshly in red, it seems likely that students interpret feedback in red as more critical as well. Indeed, given what we know about how red affects instructors' grading behaviors, students who judge feedback given in red to be more negative may be quite accurate.

Only one study that we know of has explored the question of how students evaluate instructors based on an instructor's feedback in various colors. Dukes and Albanesi (2013) had students read a student-peer's essay and an instructor's feedback in either red or aqua. They found that the instructor was evaluated as less expressive (i.e., less nice, enthusiastic, or with good rapport) when using red feedback compared with aqua feedback, but no differences were found for the level of the perception of instructor's knowledge. This research serves as suitable foundation for the current study, but we seek to extend it in a number of ways. First, we seek to increase the ecological validity of their study by using an alleged actual student's and instructor's comments in the context of a departmental assessment—Dukes and Albanesi presented participants with a hypothetical student's essay and hypothetical instructor's comments. Second, we present written feedback to students in actual handwriting—their study presented feedback written in a computer font that was meant to approximate handwriting. Third, we assess students' attributions of the student-author of the essay as well as attributions regarding the instructor. We believe it possible that attributions of one are related to the other. Finally, Dukes and Albanesi concluded that "one potential trigger in the emotional experience of receiving criticism, the color of the grading pen, affects perceptions of teaching" (p. 99). However, the authors made this claim without demonstrating whether red affected perceptions of the instructor via emotional response. We seek to model emotion as intervening between the effect of color and interpersonal attributions.

Research supports the idea that negative interpersonal attributions result from exposure to feedback in red as compared with feedback in other colors. In the present research, we seek to assess whether such attributions result from the arousal of negative emotion elicited by the color of ink used by instructors to write feedback on a paper.

We propose the following hypotheses. Compared with feedback in green or black:

H1: Feedback in red causes recipients of the feedback to be viewed as experiencing more negative emotion.

H2: Feedback in red causes a student-author's work to be perceived to be of lower quality.

H3: Feedback in red causes a student-author to be perceived to be of lower ability.

H4: Feedback in red causes an instructor who evaluated a student's academic performance to be perceived to be more critical.

In the course of investigating these four hypotheses, we will first evaluate whether instructional feedback, regardless of color, has an effect on the perception of the quality of the person responsible for the material that is being evaluated. Thus, we ask the following research question:

RQ1: Regardless of color, does the presence of feedback affect the perception of work quality?

A number of academic achievement studies have found no difference between chromatic colors like green or blue and achromatic hues like black, white, or gray (Elliot et al., 2007; Houtman & Notebaert, 2013; Moller et al., 2009; Shavit et al., 2013; Tanaka & Tokuno, 2011). We have no theoretical reason to suspect that comments in green differ on the variables of interest from comments in black; we will treat this matter as a research question:

RQ2: Does feedback in black differ from feedback in green in terms of the assessment of work, the assessment of the student who provided the work, or the assessment of the instructor who provided the feedback for the work?

In recognition of Weiner's (1986) view that negative emotional responses precede interpersonal causal ascriptions, we seek to identify if and how emotion generates responses to an instructor's colored feedback by causing attributions regarding the student and the instructor. Indeed, red is associated with negative emotions like anger (Fetterman, Robinson, Gordon, & Elliot, 2011; Wiedermann et al., 2015) and unhappiness (Brooker & Franklin, 2016). Negative emotion is expected to mediate the relationship between the color of the feedback provided by an instructor and the attributions toward the academic work, the student, and the instructor. Further, we expect that attributions toward the academic work precede attributions toward the instructor and student. Finally, we predict that interpersonal attributions toward the student and instructor are nonrecursive: Assessments of student ability influence observers' assessments of instructor competence and vice versa. We test other plausible models to distinguish whether the hypothesized attribution model provides a superior explanation for color's effect on attributions compared with competing explanations. We propose:

H5: Emotion mediates the effect of feedback color on assessments of the student's work, the student-author, and the instructor who provided the evaluation.

Method

Participants

Participants were 180 students at a Southern university from a variety of communication courses who received a small amount of extra course credit. Most were female (55%). Participants were between 18 and 27 years old (M = 19.86, SD = 1.56, Mdn = 20). Participants self-reported their race as White (58%) and African American (11%), with fewer than 5% in any other listed group.

Procedure

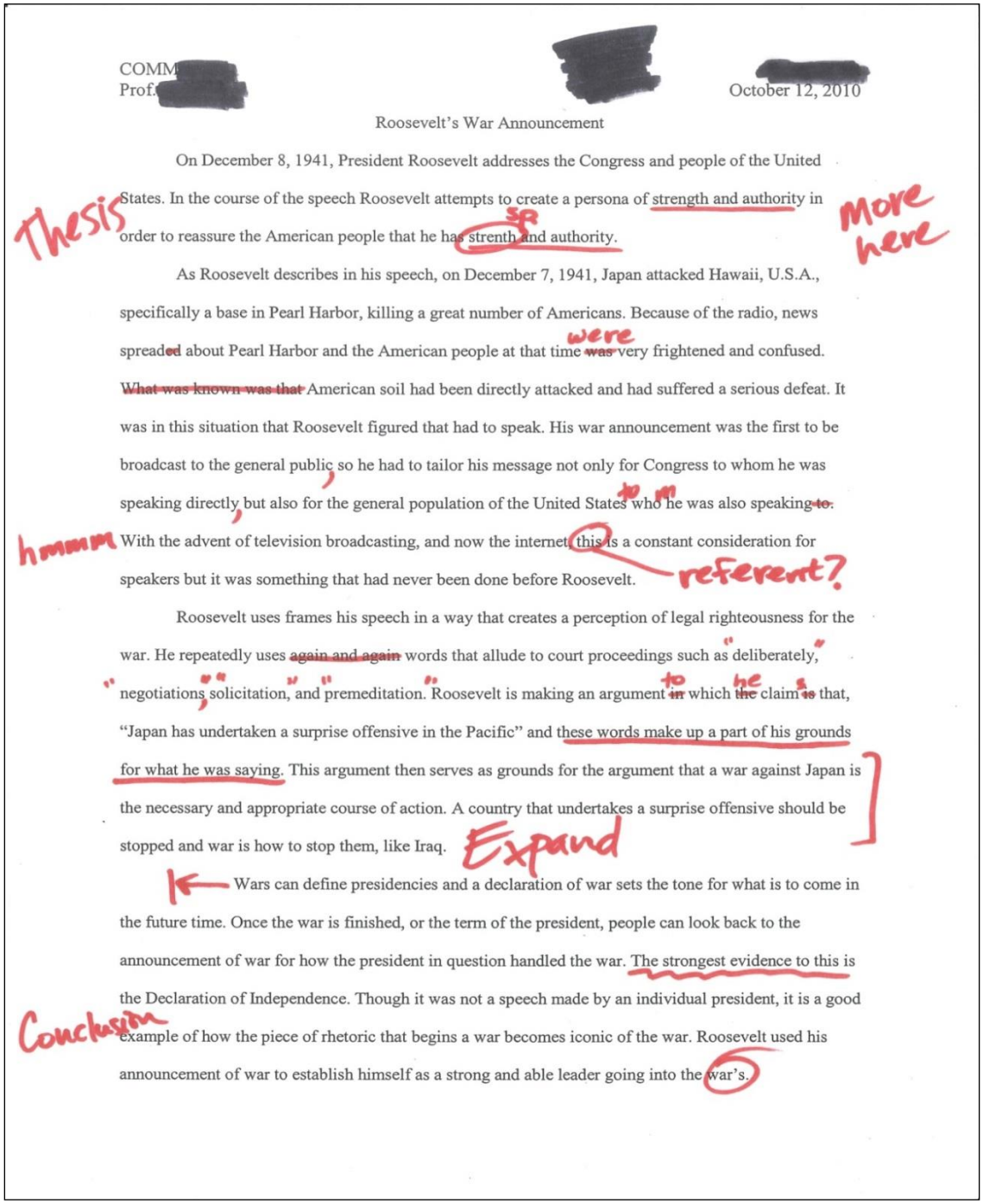
Participants were recruited via an online research management system. After agreeing to participate, they were directed to an online survey. Participants were informed that the alleged purpose of the study was a departmental review of students' and instructors' academic performance. After giving informed consent, participants were randomly assigned to a color condition. They were directed to review a scanned copy of an essay allegedly written by an undergraduate student for a communication course at their university that was evaluated by the course instructor. After reviewing the essay, the participants answered questions about their assessment of the essay, the student who allegedly wrote it, and the instructor who allegedly graded it. Participants then were debriefed as to the purpose of the study.

Experimental Stimuli and Design

Participants were randomly assigned to one of four levels of the independent variable, the color of the ink allegedly used by an instructor to provide feedback on a student's essay (i.e., feedback in red, green, or black, or no feedback at all).

An undergraduate student's essay previously written for a communication course was adapted to serve as the assignment reviewed by participants. The researchers shortened the essay and added spelling, grammatical, and logical errors. Using a fine-point marker, feedback was written in the margins, and other corrective markings were added within the text. This feedback was designed to lack obvious evaluative connotations. The student's name, professor's name, course number, and alleged grade were excised by overwriting them with black ink to give the impression that identifying information was removed to preserve student and instructor anonymity.

The essay was digitally scanned both before (to serve as the no-feedback control) and after feedback was written with red ink on the paper. A version of this essay with red feedback appears in Figure 1. The color of the feedback was digitally altered with Adobe Photoshop editing software. The red ink color was manipulated to appear as black by selecting the grayscale option and adjusting the saturation. To create the green ink condition, the following adjustments were made to the three characteristics of color in the "Reds" setting: hue +136 units, saturation −10 units, and lightness −38 units.



COMM [redacted]
Prof. [redacted] [redacted] October 12, 2010

Roosevelt's War Announcement

On December 8, 1941, President Roosevelt addresses the Congress and people of the United States. In the course of the speech Roosevelt attempts to create a persona of strength and authority in order to reassure the American people that he has strenth and authority.

As Roosevelt describes in his speech, on December 7, 1941, Japan attacked Hawaii, U.S.A., specifically a base in Pearl Harbor, killing a great number of Americans. Because of the radio, news spreaded about Pearl Harbor and the American people at that time was very frightened and confused. What was known was that American soil had been directly attacked and had suffered a serious defeat. It was in this situation that Roosevelt figured that had to speak. His war announcement was the first to be broadcast to the general public, so he had to tailor his message not only for Congress to whom he was speaking directly, but also for the general population of the United States who he was also speaking to. With the advent of television broadcasting, and now the internet, this is a constant consideration for speakers but it was something that had never been done before Roosevelt.

Roosevelt uses frames his speech in a way that creates a perception of legal righteousness for the war. He repeatedly uses again and again words that allude to court proceedings such as deliberately, negotiations, solicitation, and premeditation. Roosevelt is making an argument in which the claim is that, "Japan has undertaken a surprise offensive in the Pacific" and these words make up a part of his grounds for what he was saying. This argument then serves as grounds for the argument that a war against Japan is the necessary and appropriate course of action. A country that undertakes a surprise offensive should be stopped and war is how to stop them, like Iraq.

Wars can define presidencies and a declaration of war sets the tone for what is to come in the future time. Once the war is finished, or the term of the president, people can look back to the announcement of war for how the president in question handled the war. The strongest evidence to this is the Declaration of Independence. Though it was not a speech made by an individual president, it is a good example of how the piece of rhetoric that begins a war becomes iconic of the war. Roosevelt used his announcement of war to establish himself as a strong and able leader going into the war's.

Figure 1. Essay with red feedback, used as the stimulus for participant evaluation.

Measures

The variables that exhibited significant skewness were transformed to meet the assumptions of the general linear model. These transformations reduced skewness, and the transformed variables were used in the analyses that follow. When variables were transformed, both their transformed and original descriptive statistics are reported here, whereas statistics reported in the Results section reflect the transformed variables.

We measured emotional response in a number of ways. Participants answered questions about the emotions and emotionally laden impressions they imagined they would feel if they were the student receiving this instructor's feedback on the assignment. A modified version of the scale used by Dillard, Plotnick, Godbold, Freimuth, and Edgar (1996) consisted of two items (1 = *none of this feeling*, 7 = *a great deal of this feeling*) used to assess each emotional impression. For each emotion, the relevant two items were averaged. These means were *fear* (*afraid* and *scared*, *M* = 3.81, *SD* = 1.67, α = .75), *anger* (*angry* and *irritated*, *M* = 3.94, *SD* = 1.44, α = .70), *worry* (*anxious* and *worried*, *M* = 4.61, *SD* = 1.52, α = .75), *sadness* (*sad* and *happy,* reverse coded, *M* = 5.29, *SD* = 1.27, α = .74), *surprise* (*astonished* and *surprised*, *M* = 3.34, *SD* = 1.43, α = .77), *hope* (*encouraged* and *hopeful*, *M* = 2.51, *SD* = 1.22, α = .74), *shame* (*ashamed* and *guilty*, *M* = 4.11, *SD* = 1.60, α = .69), and *stupidity* (*stupid* and *smart,* reverse coded, *M* = 5.14, *SD* = 1.31, α = .58). A principal components analysis (PCA) indicated that the eight emotion composites reflected two dimensions accounting for 70% of total variance. The first component represented *overall negative emotion* (eigenvalue = 4.10, accounting for 51% of the explained variance), and the second represented *overall positive emotion* (eigenvalue = 1.52, accounting for 19% of the explained variance). Because only one emotion (i.e., hope) strongly loaded on the second component, which accounted for a relatively small proportion of the total variance of the emotion composites, all 16 emotion items—two indicators for each of eight distinct emotions—were used as indicators for a single latent negative emotion construct for the structural equation model. We expected the hope indicator to have a negative loading on the latent emotion variable.

Participants assessed the essay in a number of ways. On 0 to 100 scales, participants indicated the *grade they would assign* to the essay as a percentage (*M* = 72.16, *SD* = 9.53) as well as the *grade they estimated that the instructor gave* to the essay as a percentage (*M* = 72.92, *SD* = 7.18). Participants reported their estimate of the *number of errors* in the essay (*M* = 16.93, *SD* = 10.34; transformed: *M* = 1.97, *SD* = 0.29). They also estimated the *number of instructor comments that were negative* (*M* = 7.97, *SD* = 5.36; transformed: *M* = 2.64, *SD* = 1.00), *positive* (*M* = 1.90, *SD* = 1.95; transformed: *M* = 1.06, *SD* = 0.88), and *neutral* (*M* = 4.03, *SD* = 4.15; transformed: *M* = 1.74, *SD* = 1.01). The sum of these estimates form the *total number of instructor comments* (*M* = 13.86, *SD* = 6.67; transformed sum: *M* = 3.61, *SD* = 0.90).[2] Participants also responded to two items indicating the *amount of comments* the instructor gave on a 1 to 7 scale (*few* to *many* and *a little* to *a lot*), which were averaged to create a single score (*M* = 5.15, *SD* = 1.37, α = .94).

---

[2] The *number of errors* was raised to the ¼ power, and the *number of negative, positive, neutral*, and *total comments* were each raised to the ½ power.

Essay engagement was assessed with a modified version of a message effectiveness scale (Mitchell, Brown, Morris-Villagran, & Villagran, 2001). Participants rated their agreement (1 = *strongly disagree*, 7 = *strongly agree*) on five items about whether the essay was *interesting*, *enjoyable*, *logical*, *objective*, and whether *it held their attention* (M = 3.30, SD = 1.04, α = .84). *Essay quality* was assessed with four 7-point semantic differential items modeled after general attitude measures (e.g., McCroskey & Richmond, 1989) about whether the student's work was *bad* versus *good*, *weak* versus *strong*, *foolish* versus *wise*, and *hard to understand* versus *easy to understand* (M = 3.07, SD = 1.13, α = .86).

To assess the ability of the alleged author of the essay, participants indicated their agreement with four items modeled after McCroskey and Teven's (1999) competence measure. Participants rated their agreement (1 = *strongly disagree*, 7 = *strongly agree*) with statements about the student seeming *intelligent, to do well in college, capable*, and *logical* (M = 3.57, SD = 1.11, α = .88).

Finally, participants rated their agreement with eight descriptors of the instructor modeled after McCroskey and Teven (1999): Participants rated their agreement (1 = *strongly disagree*, 7 = *strongly agree*) with statements about the instructor as *negative, fair, critical, severe, careless, wrong, illogical*, and *subjective*. A PCA indicated that two components accounted for 60% of the variance. The first component (eigenvalue = 3.10, accounting for 39% of the variance) represented the *competence* of the instructor. The second component (eigenvalue = 1.72, accounting for 21% of the variance) represented the *severity* of the instructor. The two factors align with previous research showing differences in instrumental and expressive assessments of instructors in response to the color used to assess students' work (Dukes & Albanesi, 2013). These two component scores were used for group mean comparisons. Two latent constructs were used in the structural equations, with each latent construct consisting of the indicators that loaded most strongly on their relative component. The items *fair*, *careless*, *wrong*, *illogical*, and *subjective* formed the indicators of *instructor competence*, whereas the items *severe*, *critical*, and *negative* formed the indicators of *instructor severity*.

Structural Equation Modeling

Structural equation modeling was used to assess relationships among the latent variables and compare the fit of different models. Seven latent constructs, each with multiple indicators, were modeled to represent emotion (16 indicators), essay quality (four indicators), essay engagement (five indicators), essay grade (two indicators), attributions of student ability (four indicators), attributions of instructor competence (five indicators), and instructor severity (three indicators). A dummy code was assigned to the experimental manipulation of color (0 = *green or black* and 1 = *red*), and the variable's measurement error variance was fixed to zero. Measured items served as indicators for the dependent latent variables, and a single loading for each latent construct was fixed to 1.00 to provide a metric. Finally, the measurement errors of three indicators that referenced whether the object was viewed as *logical* (i.e., one for essay quality, one for student ability, and one for instructor competence) were allowed to covary due to the expectation that that a general method effect would cause these indicators to covary.

The covariance matrix based on the observed indicators was analyzed. Full information maximum likelihood estimation was used. Different structures for the seven constructs tested various processes of

color's effect, which were compared according to the models' parsimonious fit statistics. Four structural models were tested to assess competing explanations for the process by which the feedback color affected perceptions and emotional responses associated with receiving feedback on a graded assignment. The seven latent variables were divided into four panels of latent constructs: negative emotion, assessments of the essay (consisting of essay quality, essay engagement, and essay grade), attributions made about the student (student ability), and attributions made about the instructor (consisting of instructor competence and instructor severity).

Results

Manipulation Fidelity and Induction Checks

After all the items assessing the essay, the student-author, and the instructor were completed, participants were asked if they were colorblind. Those who responded affirmatively then indicated on a checklist the colors that they were unable to distinguish. No participants who indicated an inability to see red or to confuse red with another color had been randomly assigned to the red condition, and no participants who indicated an inability to see green or to confuse green with another color had been randomly assigned to the green condition.

A chi-square test was conducted to assess whether participants exposed to red, green, or black feedback correctly reported the color in which the feedback was written. Results indicated an overwhelming percentage of participants correctly reported the color of feedback to which they were exposed, $\chi^2(1, N = 140) = 97.88$, $p < .001$, $\varphi^2 = .70$. Thus, the manipulation was deemed to be successful.

Does the Presence of Feedback Matter?

To answer *RQ1* as to whether the instructor's feedback affected perceptions of essay quality, independent-sample *t* tests were conducted whereby participants in the no-feedback condition were compared with all participants whose essays contained instructional feedback, regardless of ink color. First, participants exposed to feedback estimated significantly more errors in the essay ($M = 2.01$, $SD = 0.27$) compared with those who read the essay without feedback ($M = 1.81$, $SD = 0.30$), $t(176) = 4.03$, $p < .001$, $\eta^2 = .08$. Second, participants in the feedback conditions assigned the essay significantly lower grades ($M = 71.22$, $SD = 9.22$) than those in the no-feedback condition ($M = 75.40$, $SD = 10.00$), $t(176) = -2.47$, $p < .05$, $\eta^2 = .03$. Third, participants in the feedback conditions estimated that the instructor gave the essay a lower grade ($M = 72.09$, $SD = 6.87$) than the participants in the no-feedback condition ($M = 75.78$, $SD = 7.56$), $t(176) = -2.92$, $p < .01$, $\eta^2 = .05$. Thus, the presence of instructional feedback negatively affected perceptions of the quality of the student's academic work.

Does the Particular Color of the Feedback Matter?

We hypothesized that feedback in red, as compared with green or black, would negatively affect perceptions of the student's essay, the student, and the instructor. Green and black were not expected to be different, so predictions emphasized how these colors independently related to red. Independent-sample *t* tests were conducted to assess whether the red condition significantly differed from the black condition and from the green condition on the dependent variables. Table 1 reports these results.

Table 1. Means (Standard Deviations) of Responses to Essay by Color of Feedback.

	Color of instructor comments			**Red vs. Green**		**Red vs. Black**	
Variable	Red	Green	Black	*t*	Cohen's *d*	*t*	Cohen's *d*
Affect							
Fear	3.93 (1.62)	3.67 (1.76)	3.48 (1.72)	−0.73	0.15	−1.22	0.27
Anger	4.49 (1.22)	3.43 (1.58)	3.89 (1.41)	−3.59**	0.76	−2.08*	0.46
Worry	5.03 (1.20)	4.57 (1.82)	4.40 (1.50)	−1.45	0.30	−2.14*	0.47
Sadness	5.85 (1.00)	5.10 (1.26)	5.24 (1.22)	−3.16**	0.66	−2.50*	0.55
Surprise	3.76 (1.30)	3.21 (1.61)	3.21 (1.37)	−1.83	0.38	−1.91	0.41
Stupidity	5.67 (1.13)	5.13 (1.34)	5.03 (1.14)	−2.12*	0.44	−2.61*	0.56
Hope	2.08 (0.93)	2.56 (1.30)	2.67 (1.20)	2.01*	−0.43	2.52*	−0.55
Shame	4.67 (1.48)	3.93 (1.63)	3.90 (1.64)	−2.30*	0.48	−2.26*	0.49
Overall negative affect	0.43 (0.84)	−0.20 (1.07)	−0.11 (0.96)	−3.06**	0.66	−2.62*	0.60
Overall positive affect	−0.19 (0.93)	−0.09 (1.06)	0.03 (0.90)	0.51	−0.10	1.08	−0.24
Essay assessment							
Instructor grade	70.16 (6.54)	72.30 (7.43)	73.85 (6.03)	1.48	−0.31	2.69**	−0.60
Participant grade	69.97 (7.55)	70.31 (10.52)	73.73 (8.68)	0.18	−0.04	2.12*	−0.46
Essay	3.07	3.27	3.40	0.89	−0.19	1.50	−0.34

engage	(1.02)	(1.13)	(0.95)				
Essay quality	2.67	3.07	3.35	1.79	−0.37	3.01**	−0.66
	(1.03)	(1.12)	(1.03)				
Number of errors$_a$	2.04	2.03	1.96	−0.30	0.04	−1.69	0.38
	(0.19)	(0.33)	(0.23)				
Negative comments$_a$	2.97	2.66	2.54	−1.52	0.31	−1.95	0.43
	(0.93)	(1.06)	(1.08)				
Positive comments$_a$	0.93	0.73	1.12	−1.17	0.23	0.94	−0.21
	(0.90)	(0.82)	(0.90)				
Neutral comments$_a$	2.06	1.70	1.91	−1.72	0.35	−0.70	0.15
	(1.13)	(0.94)	(0.81)				
Total comments$_a$	4.03	3.50	3.64	−2.96**	0.63	−2.35*	0.53
	(0.74)	(0.94)	(0.72)				
Amount of comments	5.91	5.01	4.85	−3.57**	0.74	−3.89***	0.86
	(1.04)	(1.39)	(1.42)				
Student assessment							
Student's ability	3.16	3.53	3.85	1.62	−0.34	3.03**	−0.66
	(1.03)	(1.15)	(1.06)				
Instructor assessment							
Competence	−0.06	−0.06	0.14	−0.03	0.00	0.88	0.20
	(1.10)	(1.01)	(0.89)				
Severity	−0.23	0.05	0.18	1.35	−0.28	2.00*	−0.43
	(0.91)	(1.07)	(0.98)				

Note. Variables with subscript *a* were transformed. The *degrees of freedom* was 95 for the red versus green test and 82 for the red versus black test. *p < .05. **p < .01. ***p < .001 , two-tailed.

Feedback in red, in comparison to green, caused participants to imagine being more angry, more sad, more stupid, less hopeful, and more shamed if they were the student-author. The participants imagined feeling more negative emotion overall. Feedback in red, in comparison to black, caused participants to imagine being more angry, more worried, more sad, more stupid, less hopeful, and more shamed if they were the student-author. The participants imagined feeling more negative emotion overall. These findings supported *H1*: Compared with feedback in green or black, feedback from an instructor that was in red caused participants to experience more negative emotions when imagining themselves as the student receiving the feedback.

Feedback in red, as compared with feedback in green, resulted in the perception that the instructor gave the essay more total comments overall.[3] Feedback in red, as compared with feedback in black, resulted in the perception that the essay would receive a lower grade from the instructor, and the

[3] Overall comments were assessed by both *total comments* (i.e., the sum of positive, negative, and neutral comments) as well as *amount of comments* (as measured by the composite of two 7-point items ranging from *few* to *many* and *a little* to *a lot*).

participant also assigned it a lower grade as well; that the essay was of lower quality; that the essay was of lower quality; and that the instructor gave the essay more comments overall. Particularly for the red–black comparison, these results provided support for *H2*: Feedback from an instructor that was in red caused the student's work to be perceived to be of lower quality.

Red, in comparison to black, caused the student-author to be viewed as less capable. These findings partially supported *H3*: Compared with feedback in black, feedback from an instructor that was in red caused the student who wrote the essay to be perceived to have lower ability. However, the green–red comparison, while trending in the expected direction, was not significant.

Feedback in red, as compared with feedback in green, caused no difference in the perception of the instructor as either more incompetent or more severe. However, feedback in red, as compared with feedback in black, led to the perception of the instructor being significantly less severe. These findings did not support *H4*: Compared with feedback in green or black, feedback from an instructor that was in red did not cause the instructor who graded the essay to be perceived to be more critical. Red, compared with black, actually appeared to cause the instructor to be viewed as less critical.

RQ2 asked whether the feedback in green differed from the feedback in black on our variables of interest. To answer this question, an analysis of variance with the three color conditions entered as levels of the independent variable was conducted for a post hoc examination of the green versus the black conditions. Bonferroni post hoc tests indicated that, although the differences between the red condition and the black and green conditions were consistent with the results reported above, there were no significant differences between the green and black conditions on any outcomes.

Model Testing

The fifth hypothesis predicted that negative emotion would mediate the effect of color on perceptions regarding the essay, the student, and the instructor. Feedback color, the experimental variable, was treated as a dummy variable, with *red* = 1 and *green* or *black* = 0. A number of competing causal models using the experimental variable and seven latent constructs were assessed and compared. The dependent variables were divided into four panels (consisting of a total of seven latent constructs): (1) latent negative emotion, (2) assessments of the essay (consisting of latent essay quality, latent essay engagement, and latent essay grade), (3) attributions made about the student (consisting of latent student ability), and (4) attributions made about the instructor (consisting of latent instructor competence and latent instructor severity). The different causal structures with which these constructs were modeled are detailed below.

A confirmatory factor analysis of the measurement model representing the aforementioned latent constructs resulted in acceptable model fit, $\chi^2(434, N = 138) = 760.44$, $p < .001$, $\chi^2/df = 1.75$, RMSEA = .07, 90% CI [.065, .082], CFI = .95, SRMR = .09.

The first causal model treated all dependent variables as equally endogenous (feedback color → emotion, essay assessments, student attributions, and instructor attributions). This model did not attempt

to specify the causal ordering among the endogenous variables. The errors of prediction of all seven endogenous latent constructs were allowed to covary. Thus, this model estimated the same parameters as were predicted in the measurement model and therefore results in identical acceptable fit statistics, $\chi^2(434, N = 138) = 760.44$, $p < .001$, $\chi^2/df = 1.75$, RMSEA = .07, 90% CI [.065, .082], CFI = .95, SRMR = .09.

The second model ordered the latent variables to correspond to their sequence on the questionnaire. This model treated potential mediation as a methodological artifact of the data collection process. The idea was that responses to later items were assumed to reflect the effect due to the prior assessments. The survey was originally arranged according to the specificity of the information being assessed on each scale, with items assessing the essay preceding the items assessing the student, the emotional responses, and finally, the instructor. Thus, a serial mediation model was constructed (feedback color → essay assessments → student attributions → emotion → instructor attributions); the error of prediction terms among three latent constructs within the essay assessments panel and two latent constructs within the instructor attributions panel were allowed to covary. This model had marginally acceptable fit, $\chi^2(449, N = 138) = 818.52$, $p < .001$, $\chi^2/df = 1.82$, RMSEA = .08, 90% CI [.069, .086], CFI = .95, SRMR = .13.

The third model accounted for the basic theoretical relationship suggested by attribution theory and color-in-context theory, whereby negative emotion mediates the relationship between color and perceptions. A single-stage mediation model was constructed, with negative emotion intervening between the feedback color manipulation on the other perceptual assessments (feedback color → emotion → essay assessments, student attributions, and instructor attributions); the error of prediction terms of the six latent constructs in the final panel were allowed to covary. This model had acceptable fit, $\chi^2(440, N = 138) = 765.44$, $p < .001$, $\chi^2/df = 1.74$, RMSEA = .07, 90% CI [.065, .082], CFI = .95, SRMR = .09.

Finally, the fourth model, which was also consistent with attribution theory and color-in-context theory but also accounted for the observers' attribution effect, represented a more complex relationship between essay assessments and student and instructor perceptions. The theoretical notion was that attributions toward the essay were directly contingent on negative emotion, but attributions toward the student and instructor were causally subsequent to essay assessments and, in addition, were assumed to be nonrecursively related. The latter assumption recognizes that attributions toward some people affect attributions toward others. In this model, emotion mediates the influence of feedback color on essay assessments. All essay assessments (i.e., essay engagement, essay quality, and essay grade) then predict student ability. Of the constructs in the essay assessment panel, only essay grade was allowed to predict instructor attributions (i.e., both severity and competence), which makes sense because instructors are directly tied to student grades, but only indirectly tied to essay quality and engagement via student ability. Finally, nonrecursive paths were allowed between student attributions and instructor attributions because of the possibility that interpersonal attributions regarding different people (the student and the instructor) affect each other and were otherwise unaccounted for. Error of prediction terms were also allowed to covary within the three assessment latent constructs and within the two instructor attribution latent constructs. This model had acceptable fit, $\chi^2(445, N = 138) = 773.57$, $p < .001$, , $\chi^2/df = 1.74$, RMSEA = .07, 90% CI [.065, .082], CFI = .95, SRMR = .10.

Model Comparison and Selection

The Akaike information criterion (AIC; Akaike, 1974) and Bayesian information criterion (BIC; Schwarz, 1978) were used to assess the success of each structural model in fitting the data. AIC and BIC values allow for comparison of nonnested models, with the latter imposing a more severe penalty for model complexity. Burnham and Anderson (2002) indicated that AIC differences greater than 3.0 provide positive evidence of model nonequivalency, with the lower value indicating the superior model. Kass and Raftery (1995) noted that BIC differences greater than 6.0 serve as strong evidence that the model with the lower value is superior. The fourth model, which tested the observer's attribution effect, had lower AIC (5,838.33) and BIC (6,081.30) values compared with the first (nonmediation) model (AIC = 5,847.20, *Model 4 difference* = 8.87; BIC = 6,122.36, *Model 4 difference* = 41.06), the second (questionnaire order-effects mediation) model (AIC = 5,875.28, *Model 4 difference* = 36.95; BIC = 6,106.54, *Model 4 difference* = 25.24), and the third (emotion-as-mediator) model (AIC = 5,840.20, *Model 4 difference* = 1.87; BIC = 6,097.79, *Model 4 difference* = 16.49). Taken together, these results provide strong evidence for the superiority of the observer's person-attribution model, and this model is displayed in Figure 2. Thus, *H5* was supported, with emotion mediating the effect of color on attributions.

Figure 2 shows the results, and the significant unstandardized relationships are reported. Red feedback significantly increased the experience of negative emotion (B = 0.70). Negative emotion significantly caused a reduction in essay grade (B = −2.39), essay quality (B = −0.34), and essay engagement (B = −0.18). Student ability was significantly affected by essay quality (B = 0.62). Essay grade did not have a significant association with student ability, instructor severity, or instructor competence. Finally, student ability caused a reduction in perceived instructor competence (B = −0.59), and instructor competence caused an increase in perceived student ability (B = 0.36), but neither causal path between student ability and instructor severity was significant. No modification indices suggested adding any paths that would significantly improve model fit.[4]

Discussion

This study assessed the effect of color used by instructors when providing feedback to students on graded assignments. Results showed that red feedback, as compared with green or black feedback, caused a negative emotional response and negative evaluative attributions toward the student-author. Model tests indicated that red indirectly led to perceptions of poorer quality work and a poorer quality student and, subsequently, a more competent instructor. This research has both theoretical and practical implications for specific instructional contexts as well as for general contexts when written messages are used.

[4] Although not represented in Figure 2, error covariances between essay engagement and essay grade (ψ = 4.69), essay engagement and essay quality (ψ = 0.96), essay quality and essay grade (ψ = 6.80), and instructor severity and instructor competence (ψ = −0.24) were all significant at the $p < .05$ level.

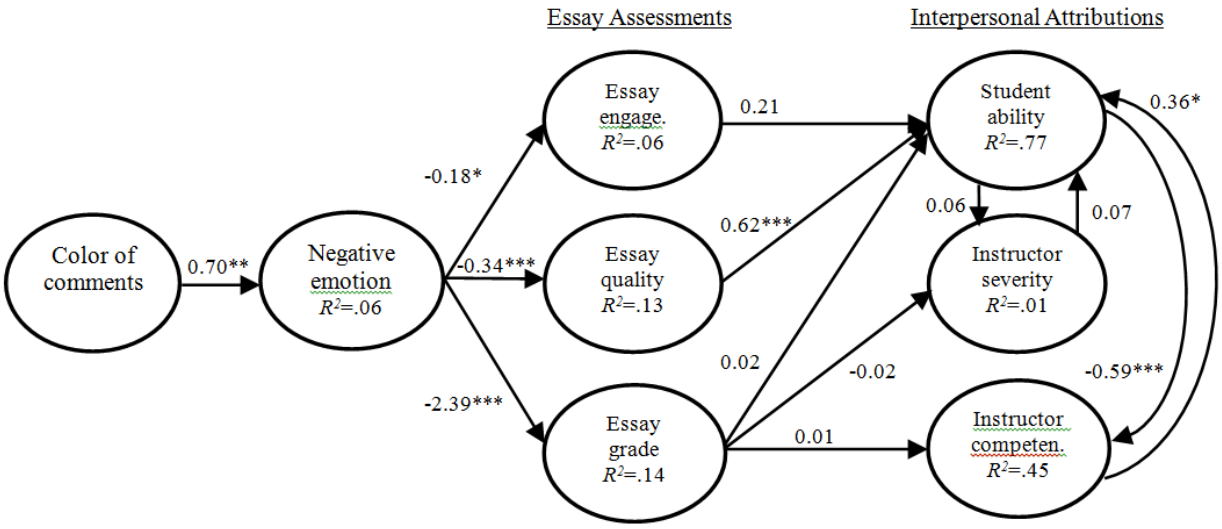


Figure 2. Structural equation model.

Coefficients are unstandardized estimates. Errors of the three latent essay assessment variables were allowed to covary, and errors of the two instructor attribution variables were allowed to covary, although these covariances do not appear in this representation of the model. Color was coded so 0 = *black or green* and 1 = *red*. *p < .05. **p < .01. ***p < .001, two-tailed.

Our findings add to the body of theoretical knowledge about how people respond to color in achievement contexts. These results agree with previous research regarding associations between red and failure in achievement settings (Moller et al., 2009). Written communication given in red appears to affect perceptions of a student's intellectual ability, as evidenced by lower grade estimates, lower perceived essay quality, and lower perceived student ability. Students appear to judge their peers to be less intelligent because of feedback in red, just as actual intelligence has been shown to suffer after exposure to the color (e.g., Elliot et al., 2007). These findings largely replicate previous research on the effects of instructor feedback written in red versus aqua (Dukes & Albanesi, 2013). However, contrary to Dukes and Albanesi, our data suggest that negative attributions were directed toward the student rather than instructor, with instructors using red ink being perceived as less severe than instructors using black ink. Further evidence for this unexpected finding—red caused students, but not instructors, to be perceived negatively—emerged in the emotion-attribution model, which we will now discuss.

In addition to influencing interpersonal attributions toward students and instructors, this study joins others (Brooker & Franklin, 2016; Fetterman et al., 2011; Wiedermann et al., 2015) to show that color significantly influenced the emotional response to the feedback. Consistent with Weiner's (1986) attributional model of motivation and emotion, emotion mediated the effect of feedback color on assessment of the essay as well as subsequent interpersonal attributions toward the student and instructor. Negative emotion negatively influenced assessments of the student's essay quality, and these

perceptions positively influenced attributions of a student's ability. In short, red elicited more negative attributions toward a student-peer through emotion and subsequent essay assessments. We also found that the effect of red on negative emotions does not translate into general negative interpersonal attributions: Our model showed that red feedback had a positive indirect effect on attributions of instructor competence. The nonrecursive student-instructor link indicates that the positive perceptions of student ability results in a decrease in the perceived competence of the instructor, whereas the perceptions of a highly competent instructor increases perceived student ability. Such a finding qualifies previous research on this issue: Whereas Dukes and Albanesi (2013) found that red diminished positive attributions toward an instructor, our model showed this to be the case in regard to instructor competence only when students were first perceived as capable.

Asking participants to make attributions about the student and the instructor provided insightful details regarding how color affects interpersonal attributions. We predicted that red would elicit greater negative attributions toward both the student and instructor. Instead, we found that red indirectly lowered attributions of the student's ability, but it indirectly increased attributions of the instructor's competence. It appears that when feedback is in red, the academic work and the student's ability are perceived as worse, but this lower quality is attributed to the student rather than to the instructor. That is, with red feedback, instructors are perceived as more competent when the quality of the work and capability of the student are perceived to be poor. In regard to black feedback, with the work being perceived as having better quality and the student as having greater aptitude, the very same feedback given by an instructor is perceived as indicating less instructor competence. These data provide support for color-in-context theory (Elliot & Maier, 2012) by emphasizing the importance of context when interpreting color's effect. Further, these findings extend color-in-context theory by identifying a boundary condition in regard to the effect of red on dependent multiperson attributions. Instead of red eliciting attributions of negativity across the board, the negativity associated with an instructor is understood in light of the negativity associated with other elements (here, the essay and student) in the situation. These results provide insight into observer attributions because students appeared to think that the instructional feedback was justified when the student was viewed as less competent due to feedback in red.

This research has a number of practical implications. The results suggest that feedback in red appears to make students think that academic work is of lower quality. However, red is also better at drawing students' attention to instructional feedback because students reported seeing more comments when they were in red. Despite the reported number of comments underestimating the actual number of these comments within the essay, red elicited more accurate estimates of the instructor's total number of comments (M = 4.03) compared with green (M = 3.50) or black (M = 3.64). These results are consistent with other findings that show that red objects are perceived to be more salient and elicit quicker responses compared with objects of other colors (Sorokowski & Szmajke, 2011). This finding suggests the need to investigate whether the attention red elicits always leads to negative outcomes. For example, although the feedback given to students in this study lacked obvious evaluative connotation, future research might assess how people respond to feedback written in red that is complimentary of a student's work. Further, students' feedback may moderate the effect of feedback color: Research shows that students high in feedback sensitivity are adversely affected by high intensity feedback (e.g., personal criticisms), whereas students low in feedback sensitivity are not as adversely affected by such messages

(Smith & King, 2004). Similarly, it is possible that red ink negatively affects students with high feedback sensitivity and not those who are less sensitive.

Ours findings suggest that educators may be in a double bind regarding the color to use to provide written feedback, given concerns for their saving their own face and their students' face. Using red may lead students to feel worse about their abilities, but they also perceive the instructor to be less severe. Using black (and, to a lesser extent, green) for feedback may lead students to feel better about their abilities, but they also perceive the instructor to be more severe. Instructors should be prepared with facework strategies in both situations, by supporting students after using red or by facilitating one's own reputation of kindness after using green.

As in all research, this study has some limitations. First, some may question the ecological validity of this study because students were asked to evaluate another student's essay rather than their own. It is possible that students respond differently after reading feedback about their own academic work compared with a peer's. For example, students attribute positive instructional feedback internally, but negative instructional feedback externally (Booth-Butterfield, 1989), so red feedback may not lead to lower perceptions of student aptitude when evaluating oneself. However, the present context of student-observers retains a level of ecologically validity: It is reasonable and consistent with our experience to assume that students share and compare their graded assignments, which result in observer assessments of classmates and instructors.

This study limited its color choices to red, green, and black, although there are other colors that teachers use to grade. It is possible that our findings are color specific, and the differences between red and green or black may not translate into feedback given in other colors, like blue and purple, or alternative shades of the colors tested here, like maroon, turquoise, or gray. Further, because the study was conducted online, participants used different computers to complete the task. Thus, it is possible that computer monitors' color settings varied such that people within the same experimental condition saw slightly different hues. That participants were randomly assigned to color conditions somewhat minimizes this issue, as systematic variations in hue between conditions were unlikely. Because of the importance of systematizing hue presentation in color research (Elliot & Maier, 2014), future studies should use greater control when exposing participants to color conditions.

This study joins a growing body of research that demonstrates the detrimental effects of red in academic achievement contexts. Written communication in red has special meaning to students. The meaning of red affects the way students respond emotionally and perceive their academic work, their abilities, and their instructors, which undoubtedly has implications for student learning. More generally, we show that the interpersonal attributions people make toward others based on the color of written messages depends initially on how they respond emotionally to such colors. Thus, this study has implications for color-induced interpersonal attributions in other achievement contexts. How nonverbal cues of color affect learning and influence perceptions of failure and success is worthy of continued investigation.

We sought to clarify the mechanism by which the color of instructional feedback influences students' perceptions. Color directly affected emotion, which then had cognitive and interpersonal effects. This research supports the notion that red does indeed prompt negative emotion, which in turn acts as a catalyst for making attributions toward others. However, as was shown in this particular context of students' perceptions of instructional feedback, not all interpersonal attributions in response to red in achievement contexts need be negative.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716–723. doi:10.1109/TAC.1974.1100705

Bertrams, A., Baumeister, R. F., Englert, C., & Furley, P. (2015). Ego depletion in color priming research: Self-control strength moderates the detrimental effect of red on cognitive test performance. *Personality and Social Psychology Bulletin*, *41*, 311–322. doi:10.1177/0146167214564968

Booth-Butterfield, M. (1989). The interpretation of classroom performance feedback: An attributional approach. *Communication Education*, *38*, 119–131. doi:10.1080/03634528909378745

Brooker, A., & Franklin, A. (2016). The effect of colour on children's cognitive performance. *British Journal of Educational Psychology, 86*, 241–255. doi:10.1111/bjep.12101

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). New York, NY: Springer-Verlag.

Dillard, J. P., Plotnick, C. A., Godbold, L. C., Freimuth, V. S., & Edgar, T. (1996). The multiple affective consequences of AIDS PSAs: Fear appeals do more than scare people. *Communication Research, 23*, 44–72. doi:10.1177/009365096023001002

Dukes, R. L., & Albanesi, H. (2013). Seeing red: Quality of an essay, color of the grading pen, and student reactions to the grading process. *Social Science Journal, 50*, 96–100. doi:10.1016/j.soscij.2012.07.005

Elliot, A. J., & Maier, M. A. (2012). Color-in-context theory. In P. Devine & A. Plant (Eds.), *Advances in experimental social psychology* (Vol. 45, pp. 61–125). New York, NY: Academic Press. doi:10.1016/B978-0-12-394286-9.00002-0

Elliot, A. J., & Maier, M. A. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology, 65*, 96–120. doi:10.1146/annurev-psych-010213-115035

Elliot, A. J., Maier, M. A., Binser, M. J., Friedman, R., & Pekrun, R. (2009). The effect of red on avoidance behavior in achievement contexts. *Personality and Social Psychology Bulletin, 35*, 365–375. doi:10.1177/0146167208328330

Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: The effect of red on performance attainment. *Journal of Experimental Psychology: General, 136*, 154–168. doi:10.1037/0096-3445.136.1.154

Elliot, A. J., & Pazda, A. D. (2012). Dressed for sex: Red as a female sexual signal in humans. *PLoS ONE, 7*, e34607. doi:10.1371/journal.pone.0034607

Feltman, R., & Elliot, A. J. (2011). The influence of red on perceptions of dominance and threat in a competitive context. *Journal of Sport & Exercise Psychology, 33*, 308–314.

Fetterman, A. K., Robinson, M. D., Gordon, R. D., & Elliot, A. J. (2011). Anger as seeing red: Perceptual sources of evidence. *Social Psychological & Personality Science*, *2*, 311–316. doi:10.1177/1948550610390051

Gnambs, T., Appel, M., & Batinic, B. (2010). Color red in Web-based knowledge testing. *Computers in Human Behavior*, *26*, 1625–1631. doi:10.1016/j.chb.2010.06.010

Hagemann, N., Strauss, B., & Leißing, J. (2008). When the referee sees red. . . . *Psychological Science, 19*, 769–771. doi:10.1111/j.1467-9280.2008.02155.x

Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: John Wiley & Sons.

Houtman, F., & Notebaert, W. (2013). Blinded by an error. *Cognition, 128*, 228–236. doi:10.1016/j.cognition.2013.04.003

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795. doi:10.2307/2291091

Knapp, M. L., Hall, J. A., & Horgan, T. G. (2014). *Nonverbal communication in human interaction* (8th ed.). Boston, MA: Wadsworth.

Little, A., & Hill, R. A. (2007). Attribution to red suggests special role in dominance signaling. *Journal of Evolutionary Psychology, 5*, 161–168. doi:10.1556/JEP.2007.1008

Maier, M. A., Elliot, A. J., & Lichtenfeld, S. (2008). Mediation of the negative effect of red on intellectual performance. *Personality and Social Psychology Bulletin, 34*, 1530–1540. doi:10.1177/0146167208323104

McCroskey, J. C., & Richmond, V. P. (1989). Bipolar scales. In P. Emmert & L. L. Barker (Eds.), *Measurement of communication behavior* (pp. 154–167). New York, NY: Longman.

McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. *Communication Monographs, 66*, 90–103. doi:10.1080/03637759909376464

Mitchell, M. M., Brown, K. M., Morris-Villagran, M., & Villagran, P. D. (2001). The effects of anger, sadness and happiness on persuasive message processing: A test of the negative state relief model. *Communication Monographs, 68*, 341–359. doi:10.1080/03637750128070

Moller, A. C., Elliot, A. J., & Maier, M. A. (2009). Basic hue-meaning associations. *Emotion, 9*, 898–902. doi:10.1037/a0017811

Rutchick, A. M., Slepian, M. L., & Ferris, B. D. (2010). The pen is mightier than the word: Object priming of evaluative standards. *European Journal of Social Psychology, 40*, 704–708. doi:10.1002/ejsp.753

Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464. doi:10.1214/aos/1176344136

Shavit, T., Rosenboim, M., & Cohen, C. (2013). Does the color of feedback affect investment decisions? *International Journal of Applied Behavioral Economics, 2*, 15–26. doi:10.4018/ijabe.2013070102

Smith, C. D., & King, P. E. (2004). Student feedback sensitivity and the efficacy of feedback interventions in public speaking performance improvement. *Communication Education*, *53*, 203–216. doi:10.1080/0363452042000265152

Sorokowski, P., & Szmajke, A. (2007). How does the "red win" effect work? The role of sportswear colour during sports competitions. *Polish Journal of Applied Psychology, 5*, 71–79.

Sorokowski, P., & Szmajke, A. (2011). The influence of the "red win" effect in sports: A hypothesis of erroneous perception of opponents dressed in red—a preliminary test. *Human Movement, 12,* 367–373. doi:10.2478/v10038-011-0043-5

Tanaka, A., & Tokuno, Y. (2011). The effect of the color red on avoidance motivation. *Social Behavior and Personality, 39*, 287–288. doi:10.2224/sbp.2011.39.2.287

Ten Velden, F. S., Baas, M., Shalvi, S., Preenen, P. T., & De Dreu, C. K. (2012). In competitive interaction displays of red increase actors' competitive approach and perceivers' withdrawal. *Journal of Experimental Social Psychology, 48*, 1205–1208. doi:10.1016/j.jesp.2012.04.004

Thorstenson, C. A. (2015). Functional equivalence of the color red and enacted avoidance behavior? *Social Psychology, 46*, 306–311. doi:10.1027/1864-9335/a000245

Turvill, W. (2014, March 19). School bans red ink and tells teachers to mark in green instead (and get pupils to respond in purple). *Daily Mail*. Retrieved from http://www.dailymail.co.uk/news/article-2584672/School-bans-red-ink-tells-teachers-mark-green-instead-pupils-respond-purple.html

Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York, NY: Springer-Verlag.

Wiedermann, D., Burt, D. M., Hill, R. A., & Barton, R. A. (2015). Red clothing increases perceived dominance, aggression and anger. *Biology Letters, 11*, 20150166. doi:10.1098/rsbl.2015.0166