Getting Beyond a Negative First Impression:

Thin-Slice Judgments of Teacher Personality

by

Fiona L. Nowlin

A Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of Bachelor of Arts

With Honors in Psychology from the

University of Michigan

2010

Advisor: Dr. Kevin F. Miller

Abstract

This study explored the temporal dynamics of personality judgments. Psychologists have identified two striking features of impression formation. We seem to be remarkably good at forming judgments of others from very "thin slices" of information, judgments that predict features such as end-of-semester course evaluations (Ambady & Rosenthal, 1993). Yet earlier research has shown that initial judgments can produce "halo effects" that affect later judgments even with the introduction of contradictory evidence. To look at the relation between these two phenomena, undergraduate participants at UM rated an elementary school teacher on 15 personality dimensions as they watched 6 short video clips. The results indicate that effects of negative first impressions dissipate quickly (within a minute) in the face of additional evidence.

Getting Beyond a Negative First Impression:

Thin-Slice Judgments of Teacher Personality

Humans have a tendency to make swift judgments of a person's character. Our initial reaction has a visceral component that often leaves a lasting impression. Thus the saying, “You never get a second chance to make a first impression.” Yet is this lasting impression based on some underlying reality, the context of the initial interaction, or our tendency to selectively use only evidence that supports our first impression? The answer to this question is important, because erroneous impressions can lead to erroneous decisions about whether and how to interact with someone.

Edward Thorndike (1927) coined the term “halo effect” to describe the way that initial impressions affect later judgments. A halo effect is defined as a tendency for initial impressions to affect later judgments even after exposure to contradictory evidence (Rosenzweig, 2007). The halo effect is a useful heuristic for harried observers, because it serves to reduce contradictory beliefs one may hold. As Rosenzweig (2007) observed, it is “a way for the mind to create a coherent and consistent picture, to reduce cognitive dissonance” (p. 50-51). Because it can be difficult to quickly measure multiple dimensions in a given situation, the halo effect can often influence one’s perception of later interactions. Halo effects may be a result of *anchoring*, in which our first impression is weighted in the initial exposure (Thorndike, 1927). In the case of first impressions of teacher personality, all subsequent interactions are then judged relative to the initial anchor.

The existence of halo effects suggests that initial impressions are often misleading. Yet other research suggests that split-second evaluations of personality can be surprisingly accurate. These small windows for decision-making are known as *thin-slice judgments*, which can be

formed very quickly. Thin-slicing explains our unconscious ability to form judgments of others from very thin slices of information, in some instances after only 100-ms of exposure (Gladwell, 2005; Willis & Todorov, 2006).

These thin-slice judgments have proven to be quite accurate in a number of situations (Ambady & Rosenthal, 1993; Carney, Colvin, & Hall, 2007; Willis & Todorov, 2006). Ambady and Rosenthal (1993) first reported that these judgments could not only be made very quickly, but also with striking accuracy. They videotaped 13 college instructors while teaching on the first day of class. Three 10-second silent clips were taken from each of the classes; the clips were 10 seconds from the beginning, middle, and end of the lessons. Nine undergraduates were paid to rate each of the clips for all of the instructors on 15 dimensions of personality. These ratings were then compared to the mean end-of-semester ratings from student evaluations of the professor by different students who were enrolled in the classes. In addition to personality, two undergraduates rated the physical attractiveness of the instructors on a five-point scale.

Overall, Ambady and Rosenthal (1993) found that students' end-of-semester teaching evaluations were highly correlated, with a mean reliability of .72, to other students' perceptions of the same teacher after only 30-seconds of exposure. This means that after much less than a minute of gathering information, participants were able to come to quite similar conclusions as students who had spent considerable in-class time with this professor. These ratings remained accurate even after adjusting for physical appearance, which may indicate that the students were influenced by very subtle non-verbal behaviors (Ambady & Rosenthal, 1993).

Yet how does one reconcile these two phenomena? Studies such as Ambady and Rosenthal's (1993) lend credence to the idea that thin-slice judgments are accurate. Though

previous experiences and expectations may cause college professors to be rated differently than elementary school teachers, there is likely to be a similar effect.

The strength of halo effects is a matter of some controversy. One study by Nisbett and Wilson (1977) found strong support for halo effects. Participants watched a video the same professor, in one condition he was likeable, respectful, and enthusiastic, in the other he was unlikeable, cold, and rigid, and then rated the professor on several different attributes. Student who saw the likeable condition of the teacher rated his appearance, mannerisms, and accent favorably, whereas the students who saw the negative condition rated these attributes as irritating. Nisbett and Wilson (1977) believed these results suggest that the holistic evaluation of a person can influence the rating of the person's specific features.

There has been some research on the accuracy of first impressions of negative and positive affect. Carney, Colvin, and Hall (2007) found that negative affect was rated accurately after only five seconds of exposure, whereas positive affect took twenty seconds or more of exposure to become significant. The authors suggest this could be related to the self-preservation in accurately judging negative situations. In addition to these results, it was also found that 60 seconds of exposure was the optimal ratio of accuracy and exposure time.

In addition to accuracy of thinly sliced judgments, there is evidence that confidence in the judgments increases when participants have additional time. Willis and Todorov (2006) found that participants who saw a photograph of a person for only 100-ms had judgment ratings that highly correlated with ratings of participants who were given 1,000-ms or unlimited time to view the photograph. The photographs were rated on

several dimensions – attractiveness, likeability, trustworthiness, competence, and aggression – and it was found that exposure time did not significantly influence ratings for any of the characteristics (Willis & Todorov, 2006). It is also interesting that in addition to confidence increasing, participants were able to differentiate between traits and as a result, characteristics became less correlated with each other.

Measuring accuracy in terms of subjective personality judgments requires a definition of this abstract concept. According to Kelley (1973) in order for a judgment to be deemed valid there must be consensus and consistency, which "suggest a means of indexing an individual's *level of information* regarding any portion of his external world" (p. 112). In order to fulfill the criteria above, multiple viewers in each condition must rate the teachers similarly (for consistency) and agree upon their overall impression of the teacher (for consensus). This operational definition allows for a working definition of what an accurate judgment might be. For the purpose of this study, accuracy was reached when there were no longer significant differences of traits between conditions.

The 1993 Ambady and Rosenthal study was the basis for the following research. If first impressions are accurate, what happens when we see uncharacteristically negative behavior as we form our first impressions? The halo effect posits that all subsequent encounters will be anchored by the initial judgment. Theoretically, thin-slice judgments should be accurate despite special circumstances of the situation or context.

We are left with two apparently contradictory predictions. We seem to be overly influenced by first impressions, and yet our judgments of others' personalities can also be surprisingly accurate based on very tiny slices of evidence. The aim of this study was to look at the dynamics of personality judgments over time in the event of a negative first impression. Are

we able to discard 10-seconds of negative information, or does it color the impressions we form of someone after a more extended exposure to their teaching? If the negative information comes after a longer period of watching the teacher, does it have the same effect on our evaluation as it would have had it come first? Answering these questions will be vital to finding a way to integrate "thin slice" research with the older "halo effect" work.

This study attempts to identify the effects of negative first impressions on teacher personality ratings. The first prediction is that, if present, halo effects should dissipate over time. That is, the difference between the means of traits in each condition will become less significant with time. It is also predicted that the negative clip of the teacher at the beginning of the video series will result in an overall more negative rating of the teacher's personality traits than if the negative video clip, or stimulus, is seen at the end of the video series. These findings are expected because participants are likely to change their opinions based on the availability of new information. By incorporating the halo effect into a thin-slice study we hope to distinguish between the accuracy of the two effects. If the halo effect theory is correct, then the teacher will be rated more negatively overall if the initial 10-second first impression is off-putting. However, if the thin-slice judgment theory is correct then there should not be much variation between the ratings in each condition.

This study was designed to capture the short-term effects of a negative first impression of an elementary school teacher. The thin-slice judgments were obtained with a series of stimulus videos, each followed by ratings of 15 personality dimensions. There were two conditions, neutral-first and negative-first. Participants saw either a negative or a neutral 10-second clip followed by four neutral clips that were identical in both conditions, and concluded with either the negative or neutral 10-second clips that was not initially seen. The halo effect was evaluated

by comparing the judgment scores between the negative-first and neutral-first conditions at all six points of rating.

Method

Participants

Participants were 40 undergraduate students (21 males, 19 females) at the University of Michigan enrolled in an introductory psychology course that participated in an online experiment in exchange for course credit. The results for the two sexes were entirely similar and therefore combined for analysis and presentation.

Students who started the survey but failed to complete it ($n = 3$), or completed the survey in less than 12 minutes ($n = 2$) were excluded from the sample. Each subject was asked to rate, on a scale from zero (not at all) to five (very much), how seriously they took their role in the study as a research participant. Two students who rated themselves as a three or less were also excluded from the sample.

Materials

The study was run using the online survey program Qualtrics and the video host website Vimeo. The first grade classroom video was taken from a previous study conducted in 2002 in Illinois. This particular female teacher was selected because she seemed to have an isolated potentially negative moment during the filmed lesson. The footage following this clip was edited into twenty-second, forty-second, one-minute, and two-minute clips. Another neutral 10-second clip was also cut to use as a control stimulus. This was done for each survey (see Figure 1). Pilot studies projected that there would be significant effects between the negative and neutral conditions for several dimensions of teacher personality.

There were two separate surveys, one for the negative-first condition and one for the

neutral-first condition. Each survey had six video clips for a total of four minutes and twenty seconds of classroom footage. The clips were uploaded onto Vimeo and then embedded into the Qualtrics surveys. The four neutral clips viewed in the middle were identical in each survey. Only the first and last clip varied between surveys, which are the negative or neutral 10-second stimuli.

For example, the set-up of Survey Alpha was: Negative ten-second clip, twenty-second, forty-second, one-minute, two-minute, and neutral ten-second clip; and the set-up for Survey Beta was: Neutral ten-second clip, twenty-second, forty-second, one-minute, two-minute, and negative ten-second clip.

The negative clip was taken from the beginning of the class, when the teacher was explaining her expectations for the students in a relatively condescending tone. The way in which she used the word "expectation" it seemed she was threatening the first-graders. She then asked the students what her expectation for their behavior was if they have a question or a comment. The answer of "raise your hand" was followed by a seemingly rhetorical question of "Does that mean I'm going to call on you right away?" Her overall demeanor struck pilot viewers as icy and uninviting, accompanied by a guarded stance and rigid facial expression.

The neutral stimulus was the 10 seconds following the negative stimulus in the full length classroom video. The teacher began the instructional portion of her lesson, she directed the students to look at their worksheet and then to select one wooden cube used for counting. Her mood is substantially more cheery than in the negative stimulus.

Procedure

After viewing each clip the participants were instructed to rate the teacher on 15 different personality dimensions on a scale of zero (not at all) to five (very much). The teacher was rated

on the dimensions of Accepting, Active, Anxious, Attentive, Confident, Competent, Dominant, Empathetic, Enthusiastic, Honest, Likeable, Optimistic, Professional, Supportive, and Warm (see Appendix for sample scale). Subjects rated the teacher after viewing each of the six clips, for a total of six Times. Each subject was then given an opportunity to comment on anything he or she noticed after the final rating scale.

Results

In the first hypothesis, it was expected that effects of the initial negative stimuli would dissipate over time, and we were interested in how fast this would occur. As Figure 2 shows, most ratings at Time 1 (10 seconds elapsed) between the negative and neutral conditions varied significantly, but the difference in ratings disappeared completely after rating Time 3 (70 seconds elapsed). There were nine traits that exhibited significant differences between the negative and the neutral condition at Time 1 (10 seconds elapsed): Accepting, $F(1,38)=9.60$, $p < .01$, Competent, $F(1,38)=1.98$, $p < .01$, Dominant, $F(1,38)=34.09$, $p < .001$, Empathetic, $F(1,38)=13.44$, $p < .001$, Enthusiastic, $F(1,37)=10.84$, $p < .01$, Likeable, $F(1,38)=33.03$, $p < .001$, Optimistic, $F(1,38)=13.92$, $p < .001$, Supportive, $F(1,38)=17.53$, $p < .001$, and Warm, $F(1,37)=37.38$, $p < .001$.

At Time 2 (30 seconds elapsed) there were five traits with significantly different means between conditions: Competent, $F(1,38)=5.\ 56$, $p < .05$, Dominant, $F(1,38)=9.29$, $p < .01$, Likeable, $F(1,38)=5.03$, $p < .05$, Optimistic, $F(1,38)=5.88$, $p < .05$, and Professional, $F(1,38)=5.38$, $p < .05$. There were no significant differences between condition of any attributes during Time 3, Time 4, Time 5, or Time 6.

The second hypothesis was also supported by the results. It was found that the average mean for eight of the fifteen attributes were significantly different based on whether the negative

stimuli at the beginning or the end of the video series. There were significant differences between the means of the following traits: the mean of Active increased from *M*=2.75, *SD*=1.01, to *M*=3.28, *SD*=1.68, *t*(32)=-1.16, $p < .01$; the mean of Confident decreased from *M*=3.45, *SD*=.88, to *M*=3.36, *SD*=1.39, *t*(32)=-.19, $p < .05$; the mean of Competent decreased from *M*=4.0, *SD*=.92, to *M*=3.43, *SD*=1.39, *t*(32)=1.44, $p < .05$; the mean of Dominant decreased from *M*=4.8, *SD*=1.05, to *M*=4.29, *SD*=.91, *t*(32)=2.23, $p < .05$; the mean of Likeable increased from *M*=1.2, *SD*=.83, to *M*=2.0, *SD*=1.63, *t*(31)=-1.86, $p < .005$; the mean of Optimistic increased from *M*=1.65, *SD*=.93, to *M*=2.5, *SD*=1.61, *t*(32)=-1.95, $p < .005$; the mean of Supportive increased from *M*= 1.55, *SD*= .99, to *M*= 2.21, *SD*= 1.63, *t*(32)=-1.48, $p < .05$; and the mean of Warm increased from *M*= .79, *SD*= .71, to *M*= 1.77, *SD*= 1.64, *t*(30)=-2.32, $p < .001$. Figure 2 demonstrates the significant convergence of traits between conditions.

Discussion

The results supported the first hypothesis, as the effect of the negative stimulus decreased after Time 2 (30 seconds elapsed) and disappeared completely after Time 3 (70 seconds elapsed). This is significant because it suggests that there is only a brief halo effect from the initial negative condition. After Time 1, participants who viewed the negative condition rated the teacher higher on Dominant, but lower on Accepting, Competent, Empathetic, Enthusiastic, Likeable, Optimistic, Supportive, and Warm, than participants who had seen the neutral stimulus.

At Time 2 there was no longer a difference between the conditions on Accepting, Empathetic, Enthusiastic, or Supportive. There was still significant variation for the means of Competent, Likeable, Optimistic, and Warm. However, there was a significant difference between the ratings of Professional that was not evident in Time 1. As previously stated, after

Time 3 there were no longer any significant differences in attributes between conditions.

Due to the fact that significant differences were found in Time 1, fewer significant differences were found in Time 2, and no significant differences were found after Time 3 it is clear that the ratings converge to become accurate, or stable. Though a small halo effect is seen, the evanescent nature of the effect suggests that the participants were able to quickly adjust their opinion of the teacher's personality based on new information. As the exposure time to the teacher increased the level of information available altered most personality ratings. It is quite remarkable that the negativity in the first impression could be corrected in 30 seconds.

In the Ambady and Rosenthal (1993) article, the students who rated the professor in the end-of-semester evaluations were also present on the first day of class. Therefore, it is likely that their opinion included the information that the thin-slicing participants used to judge personality. However, in the case of our findings, differences between the ratings disappeared before the neutral-first participants were exposed to the negative stimuli. What is even more interesting is that after both groups had been exposed to all of the clips, no significant difference in ratings appeared after the introduction of the negative stimuli to the neutral-first condition. This attests to the strength of judgments after less than a minute of exposure. This stability in ratings is an example of how thin-slicing can be accurate after a short period of time, but confidence in the accuracy of ratings increases with the amount of exposure (Willis & Todorov, 2006). According to Carney, Colvin, and Hall (2007) the optimal accuracy-exposure time ratio is 60 seconds. In our study, the effect dissipates completely between 30 seconds and 70 seconds of exposure, which is in accordance with Carney, Colvin, and Hall (2007).

As predicted by the second hypothesis, there was an interaction between Time and Condition for the negative stimuli. Using an independent *t*-test for the ratings of the negative

stimuli between the conditions, it was possible to compare the means of each trait while distinguishing whether they were seen as a first impression, at Time 1, or a last impression, at Time 6. When seen as a last impression, students generally rated the teacher's personality more positively, compared to when the same 10-seconds was the first impression. As indicated by the *t*-tests, subjects rated the teacher as less Dominant, and more Active, Confident, Competent, Likeable, Optimistic, Supportive and Warm, if the negative stimulus was at the end of the series of clips.

These results support the conclusion that the five neutral clips of the teacher acted as a buffer against the adverse affects of the negative 10-second stimulus. For example, the teacher had a mean rating of 1.2 for Likeable when the negative clip was seen first, compared to a mean of 2.0 when the negative clip came at the end of the series, $p < .005$. Other traits that differed significantly were Active, Competent, Confident, Dominant, Optimistic, Supportive, and Warm, $p < .05$. Conversely, if the negative stimulus was seen as a first impression, at Time 1, it was rated much more negatively than if it was seen as the last of six videos, at Time 6. There were no traits that were significantly different at Time 6 but not significant at Time 1.

As the results suggest, the adverse affects of the negative first impression disappeared quickly. There were virtually no differences between the conditions after the third video. This is an important finding as it indicated that there is an initial effect of the first impression, but that it then is modified to reflect further information discordant with the first impression. As humans, we are able to adjust our perceptions based on the availability of new information. As shown in Figure 2, many of the traits converge completely by the fifth rating.

The topic of this study is one that affects everyone. Every day we judge and are judged by others. Sometimes these judgments are erroneous or based on uncharacteristic behavior. So it

is important to know whether or not first impressions of personality judgment are ephemeral and constantly being edited with the availability of new information. There are times when first impressions seem crucial, such as a job interview or a first date. Being aware of the true effects of first impressions may help reduce anxiety in stressful situations.

This study had several limitations. First of all, these results were based on a single teacher. It is quite possible for judgments of another teacher to produce different results. This could be due to the salience of the negative first impression, physical appearance, the teacher's overall negative disposition, the reaction of the students to the teacher, and a number of other confounding variables. There is a lot of variation between what is considered negative to different people. It is also important to mention the limited sample of participants, who were roughly the same age and all students at the same university. These factors could have influenced ratings because of participants' past experience or expectations of the teacher or the study itself.

Future directions for this kind of research could investigate the influence of students' perceptions of their instructor and how it impacts other factors, such as attendance, achievement, self-reported conduct, or participation. It would be interesting to see the effects of different personality characteristics on classroom dynamics as reported by the students. This information could be used to help prepare novice teachers for what students respond to best. The ultimate study would be a real-world experiment in which participants rate a confederate's behavior after a negative first impression, then a separate rating in a follow-up study.

The main conclusion of this study is that although first impressions are important, people are capable of adjusting their opinions based on new information. The results of this study suggest that a negative first impression may be a problem for no more than five minutes at most.

A popular saying asserts, "You never get a second chance to make a first impression."

Our results suggest that this first impression can be fleeting, which should be good news for everyone who has ever fallen on their face in a first class or meeting.

References

Ambady, N., & Rosenthal R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, *64*(3), 431-441.

Carney, D. R., Colvin, C. R., & Hall, J.A. (2007). A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality*, *41,* 1054-1072.

Gladwell, M. (2005). *Blink: The Guide to Thinking without Thinking*. New York: Little, Brown, and Company.

Kelley, H. H. (1973). The process of causal attribution. *American Psychologist, 28*(2), 107-128.

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*(4), 250-256.

Rosenzweig, P. (2007). *The Halo Effect…and the Eight Other Business Delusions that Deceive Managers.* New York: Simon and Schuster, Inc.

Thorndike, E. L. (1927). The law of effect. The *American Journal of Psychology, 39,* 212-222.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*(7), 592-598.

Author Note

Fiona L. Nowlin, Department of Psychology, University of Michigan.

Kevin F. Miller, Department of Educational Studies and Psychology, University of Michigan.

I owe an enormous amount of gratitude to my mentor, Dr. Kevin Miller, whom I was honored enough to worked with for three years. I am so grateful for your encouragement, support, and guidance. You have truly been an integral part of my experience at Michigan. I also owe thanks to two awesome graduate students, Scott McCann and Fernando Rodriguez, Jr.; Dr. Vilma Mesa, for all her patience while explaining statistics; a huge thank you to the wonderful faculty and staff in the Psychology Department, especially Lucy Thao and Kristi MacKenzie; and finally Lynn Chamberlain, thank you Mom – you're the greatest.

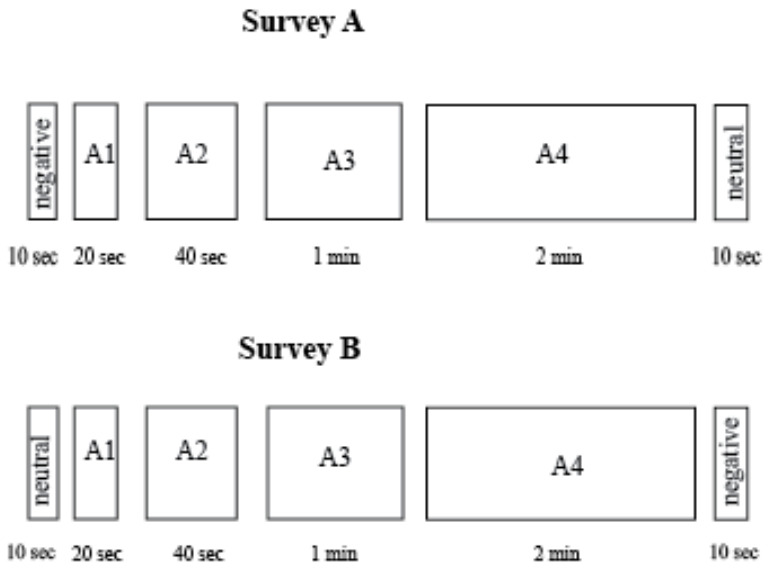

Figure 1. Chart as visual demonstration of duration and condition in each survey.

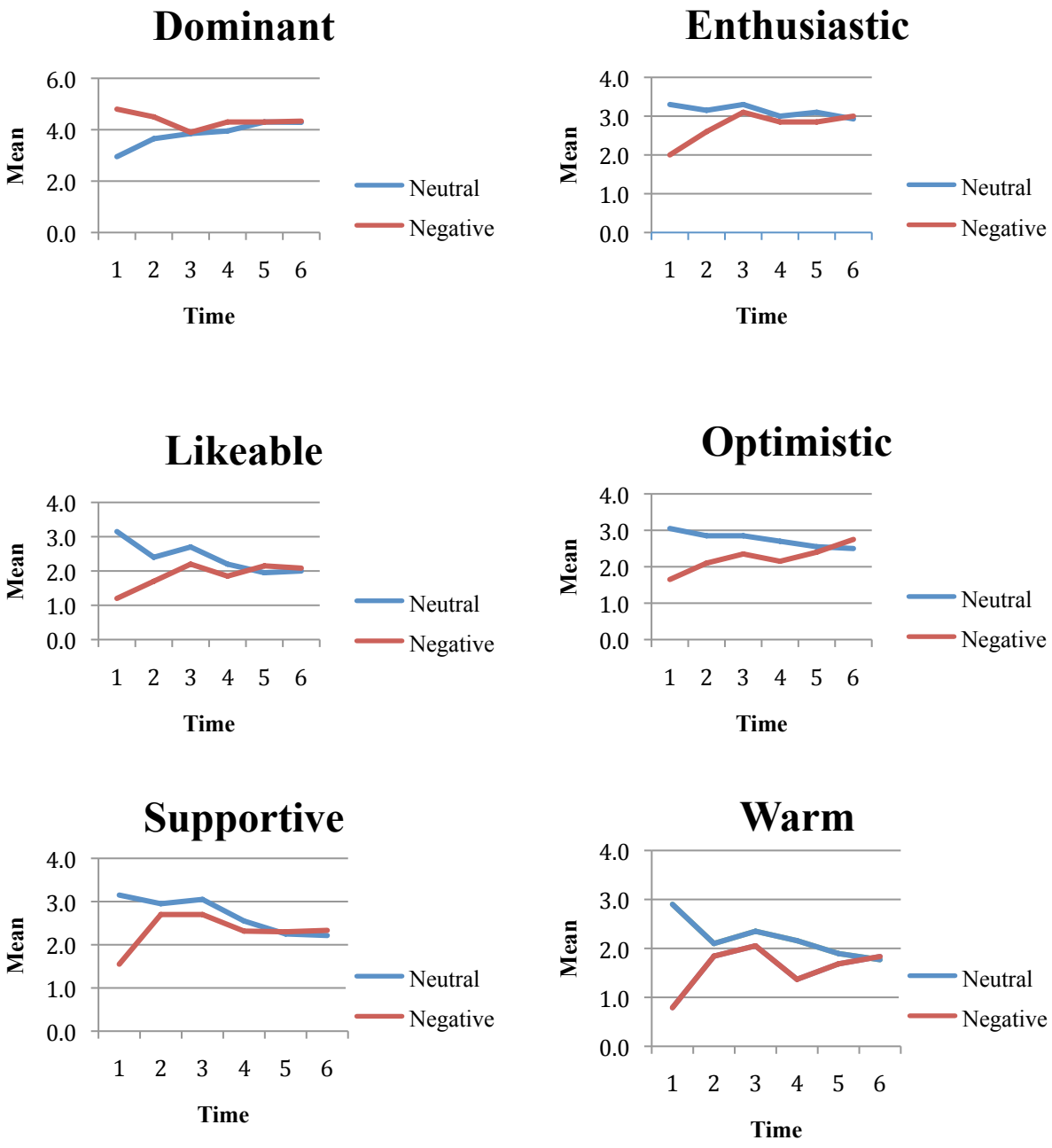

Figure 2. Graphs demonstrating significant convergence of trait ratings, $p < .05$.

Appendix

Sample Rating Scale of Teacher Personality Dimensions