10-30-2017 1:30 PM

Does Reflection Mitigate Negative Emotions Following Work Performance Feedback?

Rebecca J. Factor, *The University of Western Ontario*

Supervisor: Goffin, Richard D., *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Psychology
© Rebecca J. Factor 2017

Abstract

This study examined if reflection could mitigate negative emotion following negative work performance feedback. Initial research has found that reflection is beneficial for learning, but it has seldom been tested if reflection can mitigate negative emotion associated with negative feedback. Participants were tasked with completing open-ended questions based on a workplace training manual, and then received negative work performance feedback. Feedback was presented in either absolute terms, or relative to others' performance. Afterwards, in one condition, participants completed a reflection activity, while in another condition, participants simply completed a time filler task. Participants' emotions were then measured. Results indicated that reflection did not lead to lower scores of negative emotion. In fact, there was evidence that it resulted in higher scores of negative emotion, which held when feedback was presented in absolute terms or relative to others. Reasons for this are discussed, as well as potential future research.

Keywords: Feedback; Reflection; Performance Appraisal; Performance Management; Affect; Emotion; Feedback Acceptance; Relative Ratings.

Acknowledgements

I would like to thank my supervisor, Dr. Richard Goffin, for his support and guidance throughout this thesis project, and for many other projects we have worked on together. I've very much appreciated your advice and learned from your expertise. Thank you to Dr. Paul Tremblay for assisting me with my analyses in this project. I would also like to thank my colleagues in the I/O Psychology program, who have continually answered my questions and helped me throughout the past two years. I also want to say thank you to my family, Diane Boyes, John Dodds, David Factor, and Sherri Wolfish, for their continued support while I've been completing the program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

Does Facilitated Reflection Mitigate Negative Emotions Following Work Performance Feedback?

Feedback is imperative for both individuals and teams to develop at work (Kluger & DeNisi, 1998). Employees, however, are not always receptive to feedback, affecting the understanding and acceptance of feedback, as well as later performance.

In other words, if employees do not accept the feedback process, the effectiveness of said process is compromised (Dipboye & Pontbriand, 1981). Feedback often has little or no effect on future performance, a possible reason being that individuals receiving feedback tend to spend very little time thinking about the feedback after receiving it (Anseel, Lievens, & Schollaert, 2009). Additionally, affective reactions can interfere with an employee's reception of the feedback and later performance. This is especially probable when the feedback incorporates criticism (Belshack & Den Hartog, 2009; Kluger & DeNisi, 1998).

It is likely that feedback could actually have negative effects on work performance due to issues with reception. For example, negative feedback could produce a host of negative emotions, and a defensive response, which could subsequently interfere with improvement on the job. Recent literature suggests that certain cognitive reframing techniques, such as reflection – defined as "…the intellectual and affective activities individuals engage in to explore their experiences to reach new understandings and appreciations of those experiences" (Boud, Keogh, & Walker, 1985, p. 19) – can have certain benefits, including learning, gaining new perspective, and decreasing emotional reactions. Ultimately, this could improve the understanding and reception of feedback (Anseel et al., 2008).

Kluger and DeNisi (1996) proposed a model in which employees understand workplace feedback as being on the 'self level,' which changes focus of the feedback from being *about the task* to *about oneself*. Employees interpreting feedback at the 'self level,' which usually occurs when negative feedback is given, will experience negative emotions (Sargeant, Mann, Sinclair, Van der Vleuten, & Metsemakers, 2008; Kluger & DeNisi, 1996). Certain cognitive processes, such as specific types of reflection, can result in learning new perspectives and understandings, which could aid in combatting employees' negative emotions resulting from feedback (Sargeant et al., 2008). As Boud et al. (1985) explained, people can be exposed to a vast amount of input and have trouble assessing the information, or feel overwhelmed. When reflecting, an individual will recall an event, attempt to work through it, and set aside emotions associated with said experience in order to incorporate new perspectives. Thus, in setting emotions aside, reflection could mitigate negative reactions to feedback, ensuring feedback is both received and understood. Additionally, because an individual can gain new perspectives through reflection, they may be able to better understand feedback. This demonstrates how emotions and learning are embedded in the reflection process, and how engaging in reflection can mitigate negative emotion and enhance feedback reception.

Theoretical Frameworks Relevant to Reflection

Kolb's (2001) experiential learning theory holds that there are four stages of learning: concrete experience, reflective observation, abstract conceptualization, and active experimentation (Kolb, Boyatzis, & Mainemelis, 2001). This theory could be applied to the events associated with the feedback process. Concrete experience occurs when feedback is presented, reflective observation is equal to reflecting on feedback,

abstract conceptualization encourages learning, and applying the presented feedback would be a part of active experimentation. Overall, these four stages offer an explanation for learning from the point of a first experience, all the way to incorporating learned concepts into future behaviour. This theory presents reflection as a vital aspect of the learning and behaviour process, and provides one reason that reflection plays a role in feedback reception and later performance.

However, in order to maximize learning, reflection should be practiced in a way that results in deep level processing (Xie, Ke & Sharma, 2008). According to cognitive models, material processed at the surface does not require much thought; as such, it may not be remembered or lead to change. In contrast, deep processing is a longer process – a result of elaboration and evaluation of material (Anseel et al., 2009). Deep processing will be longer lasting, and is more likely to result in behavioural changes.

Facilitated reflection is vital for people to engage in critical thinking (Loo & Thorpe, 2002) and can encourage deep reflection; to prevent information from being processed merely at surface level. Facilitated reflection is designed to prevent individuals from reflecting freely. In order to guide the reflection and provoke a different way of thinking, they are asked to complete structured activities, i.e. reflective questions (Anseel et al., 2009). This is type of reflection is advantageous because, as Boud et al. (1985) pointed out, people often reflect on many different matters; yet, the matters they reflect on may not provoke deeper processing. In turn, what was being reflected upon may not be remembered or affect behaviour. However, if reflection is leveraged to encourage deeper processing (such as in a facilitated reflection activity), it can be more impactful for an individual, enhancing learning and modifying behaviour. In the workplace, deep

reflection could improve the understanding of feedback and influence subsequent work behaviour and performance.

In fact, Scanlan and Chernomas (1997) described a three-step model of reflection that emphasizes the importance of deep processing. The model consists of awareness, critical analysis, and learning. First, one will experience awareness, when positive or negative thoughts about an event surface. Next, critical analysis occurs. The individual will reflect on their ability to handle adversity or doubt, resulting in self-awareness and learning. The last stage, learning, should result in a realization or understanding of a new perspective (Scanlan & Chernomas, 1997). This model can be compared to the feedback process, as follows: an employee will feel either negative or positive feelings about the feedback; the employee will reflect on feedback leading to awareness; he or she will learn from the feedback and understand a new perspective. To mitigate negativity and bolster learning, the individual must experience all steps. Therefore, without deep reflection, employees simply remain 'aware' of their feelings and move forward without having gained new perspectives or realizations. This affects feedback reception and workplace behavioural change.

While it is clear that reflection plays a role in learning, less is known about its role in mitigating emotional responses. When deeply reflecting on a task or event, Boud (1985) stated that people 'set aside' emotional responses to think about the event. Yet, not all reflection theories incorporate emotion, nor do they state the benefits of reflection in terms of managing negative emotion. Furthermore, many of the current studies tend to investigate reflection and learning, but do not investigate the effects on emotion. The underlying role of reflection in emotional reactivity may be particularly important in

ensuring the effectiveness of facilitated refection. As well, emotional reactivity plays an important role related to the effectiveness of work performance feedback. Negative feedback could elicit an emotional response, which then affects learning and performance. Literature integrating emotion and reflection is scarce, especially in the field of industrial/organizational (I/O) psychology and work performance assessment. Current literature stems from the clinical and educational fields, among others. This further emphasizes the importance of examining the outcomes of reflection in the context of work performance feedback.

Research on the Benefits of Feedback and Facilitated Reflection

An experimental study in the field of computer science investigated both feedback and reflection in teams (Phielix, Prins, Kirschner, Erkens, & Jaspers, 2011). Participants were 108 high school students, randomly assigned in groups of 2, 3, or 4. Students used Radar, an online peer feedback forum, to provide team members with information pertaining to cognitive and social behaviours. Team members rated each other, from 0-4, on influence, friendliness, cooperation, reliability, productivity, and quality of contribution. Participants completed a self-evaluation, as well as an evaluation on all other group members, on Radar. Participants also used Reflektor, an online reflection tool with 4 items. An example item would be, "What is your opinion on how the group is functioning? Give arguments to support this," (Phielix et al., 2011, p. 1094). Each group member individually answered the reflection questions in Reflektor. Once completed, the answers were sent to the other group members. Two additional reflection items were answered in an online form called Co-writer, which is similar to Reflektor. This program, however, had reflection items for the whole group to answer. One of the items asked,

"What does the group think about its functioning in general? Discuss and formulate a conclusion shared by all the group members," and the other item focused on goal setting, "Set specific goals (who, what, when) to improve group performance," (Phielix et al., 2011, p. 1094). To investigate the effectiveness of the reflection tools (Reflektor and Co-writer) as well as the feedback tool (Radar), researchers manipulated when they were introduced. These tools were introduced either in the beginning – Time 1 – halfway – Time 2 – or end – Time 3. If the tools were introduced at Time 1, they were also used at Time 2 and 3. If the tools were introduced in Time 2, they were also used at Time 3. If the tools were introduced at Time 3, then they were only used at Time 3. Those participants who used the tools (either beginning or halfway) had more convergent evaluations (self versus other) at Time 3. They also had higher satisfaction, team development, positive attitudes when facing a problem, and less group conflict at Time 3. The tools did not lead to higher cognitive performance, as measured through group and individual ratings of productivity and quality of contribution. This could be due to the short period of time using the tools, or to the subjective nature of the measure. Overall, there is some evidence that receiving feedback and participating in reflection activities could be beneficial for being receptive to feedback. When the tools were used throughout the task, self-evaluation scores were similar to peer evaluation scores once the task was completed, perhaps indicating that people gained an understanding of new perspectives through reflection. Using the tools also resulted in positive benefits for the teams, further demonstrating the learning of new perspectives.

There are, however, limitations to these results. It is unknown whether the findings are due to the reflection tool in particular, or from all of the tools used in

conjunction, making the results difficult to interpret. Notably, Phielix, Prins, & Kirschner (2010) developed an original tool for Reflektor that did not have a goal setting item. Due to the null results in that study, they revised the reflection tool to include an item about future performance. Therefore, it remains somewhat unknown if the conclusion of the Phielix et al. (2011) study was due to the goal setting item, or other changes (such as sample size) between the two studies. According to the aforementioned definition of reflection, a reflection instrument should focus on the current experience, but Reflektor encourages thought on past, present, and future behaviour. There is no mention of forward planning in the definition of reflection; as such, the goal setting item should not be included in a reflection tool. This inconsistency in the results demonstrates the need to consider additional research that is focused on feedback and reflection.

Another study more directly related to the field of I/O psychology examined reflection following feedback on a task (Anseel et al., 2009). Participants were recruited through a government job training website that advertises a variety of tests people can complete to learn more about their work skills. As such, many of the participants were employees or applicants. If people volunteered to participate, they were asked to rate the effectiveness of different responses to 20 e-mails: 10 at Time 1, and 10 at Time 2. Using a computer program, participants' problem awareness, coordination, information management, and decisiveness were immediately graded. After completing the first 10 e-mails, participants were placed in a condition where they would: a) receive no feedback on their performance (they would merely be told that they are halfway done); b) only receive feedback on their performance; c) receive feedback on their performance and be asked to complete a reflection activity; or d) only be asked to complete a reflection

activity. Then, participants would respond to 10 new e-mails, which made up an alternate form of e-mails to those in Time 1. Performance in the alternate form would be compared to participants' ability in the first 10 e-mails. The reflection activity asked participants to explain (in writing) what they thought they had done well, and done poorly, on 4 different performance items. This reflection activity fits the definition of reflection because it is a guided reflection of a current experience, encouraging new thought and perspectives. Feedback was presented on the four dimensions with a rating from 1-20, alongside a short narrative explaining the results.

Participants improved their performance most if they were in the feedback and reflection condition. In the no feedback and no reflection condition, performance did not improve; in the reflection only condition, performance improved very little. In the feedback only condition, improvements were made from Time 1 to Time 2, but not to the same extent when both feedback and a reflection tool were presented. This was convincing evidence that reflection with feedback could enhance later performance. However, this was a simulation and was based on a very simple task in the workplace: writing emails. Also, participants would have been very motivated to do well, as they voluntarily participated in the task. Even so, this is an indication that reflecting after receiving feedback in the workplace could produce significant improvements in performance.

Therefore, reflection activities could enhance feedback reception through eliciting learning, but few tests exist demonstrating if reflection could deter negative emotion. In addition, there is not enough research on reflection in relation to feedback to make clear assumptions. Research that is available seems inconclusive, such as Phielix et al.'s (2011)

study in which goal setting could have been the reason for changes in group outcomes. Having said this, the theoretical evidence and proposed benefits indicate that reflection could be valuable in mitigating negative emotion (which could assist in the learning process). Further exploration of this topic would make both a research and theoretical contribution.

Present Study

We examined the effects of reflection on emotion after receiving negative feedback. We were interested in whether reflection could mitigate negative affect (NA) over and above a "time filler" control condition. We also examined whether the effects of reflection hold when two different methods are used to present the rating: an absolute feedback method, with conventional descriptive anchors, and a relative method where performance was described in comparison to others. There was no hypothesis for this aspect of the study, as previous research has not examined this. However, evaluations that use comparative language usually impact someone more negatively, as the feedback is directed to the 'self level' (Kluger & DeNisi, 1998). We wanted to investigate if reflection could still combat the NA resulting from this condition. Despite anticipating that participants who complete the reflection activity would result in lower NA scores (than those that complete the time filler task), we tested the difference in both directions, (i.e. using a two-tailed test), given that this is a new area of research and it is impossible to predict any unintended side-effects.

H$_1$: Participants who have completed a time filler task will have significantly different mean scores of NA than those participants who completed the reflection activity.

Brett and Atwater (2001) found that a discrepancy between self-ratings and others' ratings of performance predicted negative reactions to feedback. Consequently, we also assessed whether controlling for one's self-rating of performance influenced the effect of the reflection manipulation on NA. We expected that one's self-rating would be a significant covariate:

H₂: After negative work performance feedback, participants who have completed the reflection activity will have significantly different NA than will those in the control (time filler) condition, and participants' self-ratings will be a significant co-variate in this analysis.

Whether or not an employee accepts performance feedback influences the effectiveness of the feedback process (Ashford, 1986), and in turn, if employees make behavioural changes. As such, four subscales from The Feedback Acceptance Scale (Kedharnath, Garrison, & Gibbons, 2010) were included in this study: self-awareness, fairness, clarity, and intent to use. It was predicted that reflection would have an impact on the self-awareness and intent to use subscales. This is because these items are most related to outcomes of reflection; that is, gaining a new perspective and learning. Because this is a nascent area of research, we tested for differences in both directions.

H₃ₐ: Self-awareness scores will be significantly different for those in the reflection condition than for those in the control (time filler) condition.

H₃ᵦ: Intent to use scores will be significantly for those in the reflection condition than for those in the control (time filler) condition.

Method

Materials

Demographics Questionnaire. Participants completed a short demographic questionnaire containing questions on gender, age, education history, and if they have received feedback in a past job. Some of these questions were based on items asked in Derue et al.'s (2012) study. All questions are included in Appendix A.

PANAS (Watson, Lee & Tellegen, 1988). Participants completed this 20-item scale that measures both positive and negative affect to analyze emotional reactions to the feedback. The PANAS was filled out at the beginning and end of the study to see if a change in affect occurred (manipulation check), the end-of-study PANAS score served as a dependent variable. This questionnaire has high validity, as items have convergent validity with similar items, and divergent validity with items measuring a different type of affect (Watson et al., 1988). The scale was referred to as a "Feelings Scale," as it sounded less academic. The negative items used in the analysis were scared, afraid, upset, distressed, jittery, nervous, ashamed, guilty, irritable, and hostile (items #2, #4, #6, #7, #8, #11, #13, #15, #18, and #20, respectively, in Appendix D and K). Using a scale from 1-5 – from very slightly or not at all, to extremely – participants indicated how they were feeling at the present moment.

Training Manual. Participants read a "Conflict Diamonds Training Manual" in which they were asked to imagine themselves as a sales associate selling diamonds to customers. The manual provided instructions on dealing with inquiries about conflict diamonds. It was based on a manual from the World Diamond Council in Alliance with Jewellers Vigilance Committee, Jewellers of America, Diamond Dealers Club, Diamond Manufacturers and Importers Association of America, and Jewellery Information Center. After reading the manual, participants' task was to answer questions about the manual.

Most items were scenario based and open-ended so that participants were somewhat unsure of how they did. All items are included in Appendix C.

      Self-Evaluation of Your Answers to the Questions. Participants self-evaluated their overall performance on the above task using a 1 (poor) to 5 (excellent) scale. Participants were asked to rate themselves based on how accurately they answered the questions while using their own words. They were shown a scale and asked to select their score underneath the scale, as shown in Appendix D. A 5-point scale was chosen because this is what is most widely used in the workplace (3D Group, 2009).

      Bogus "Computer-Scored" Evaluation of Your Answers to the Questions. Participants received a numerical rating on their task performance after completing their self-evaluation. They were told that this rating was produced by a machine learning algorithm. In fact, all participants received 2/5, which constituted negative feedback. This score was chosen because 1/5 was thought to be either extremely upsetting for participants or seem implausible. A score of 3/5 and up did not seem negative. Therefore, 2/5 seemed like the most fitting rating for this study. In one condition, a sentence explained that 2/5 was, "better than poor but less than satisfactory" (absolute condition). In another condition, a sentence accompanying the 2/5 was, "64% of participants performed better than you" (relative condition). The percentage, 64%, was chosen for similar reasons as the 2/5 rating: because any percentage lower than 64% may seem implausible, or seriously upset participants, causing them to drop out of the study. The percentage was also low enough to be a negative score, thus corresponding to 2/5. These computer-scored evaluations are shown in Appendix E and F.

Reflection Activity. Two open-ended items were created to encourage participants to reach new conclusions by reflecting upon the experience of completing the activity and receiving feedback. Participants were asked to reflect about why there was a difference between their self-rating and the computer-scored rating. These items were influenced by the reflection activities used in previous research, such as focusing on what someone has done well or poorly (Anseel et al., 2009), and the items used in Phielix et al.'s (2011) study. From these instruments, our items were developed to create an impactful reflection activity. The reflection activity is included in Appendix H.

Time Filler Task (HEXACO-PI-R, Lee & Ashton, 2016). "Neutral" items chosen from the HEXACO-PI-R (Lee & Ashton, 2016) by four graduate students and were filled out by participants who were randomly assigned to the control condition. Items were chosen with the goal of not eliciting an emotional response. Most items were taken from the Openness to Experience and Extraversion facets. One item each from Agreeableness and Conscientiousness was also used. Potential neutral items were initially chosen from the 100-item HEXACO PI-R and looked over by four graduate students. Items repeatedly believed not to be neutral were cut from the item list. Having participants complete neutral HEXACO items was chosen as the time filler task as it would not change participants' emotional reaction. Also, it is believable that these items would be included in a research study, so participants would not be confused about the task they were completing.

Feedback Acceptance Scale (Kedharnath, Garrison, & Gibbons, 2010). Participants completed select items from 4 (out of 7) Feedback Acceptance Subscales. They were asked to think about their feedback and answer each item from 1-5 (strongly

disagree to strongly agree). The subscales in this study were self-awareness, fairness, clarity, and intent to use. The other subscales were not included because they were not applicable to the feedback presented in this study. Some of the items were adapted for the study, while others were deleted. All items included are in Appendix G.

Response Check. There were 10 items included in the response check. Based on the recommendations by Meade and Craig (2013), three items were instruction response items, and one item pertained to the quality of responses. Three items asked participants about the content of the training manual. In order to assess how real participants believed both the cover story and the negative feedback to be, two manipulation check items were included. The last question asked participants if they would like to comment on the study. This created an opportunity for participants to share their thoughts on the study, thus enhancing our understanding of the results. All items are shown in Appendix H.

Procedure

The entire study was completed online using Amazon's Mechanical Turk (MTurk) with a $2.00 USD participation incentive. Upon seeing the advertisement for this study, participants read the cover story (see Appendix A), which explained that the study was being carried out to test the effectiveness of a workplace training manual for a retail company selling diamonds.

Participants read the letter of information and the consent form (Appendix B), which also emphasized that the study was investigating the effectiveness of a workplace manual. Participants then completed the demographic questionnaire (Appendix C), as well as the PANAS (Appendix D). The PANAS measured an individual's feelings, and was referred to as "The Feelings Scale." Following this, participants read the workplace

manual and answered open-ended questions about the manual's suggestions for dealing

with customers (Appendix E). Answering these questions was the task participants based

their self-evaluations (described above) on. Participants were asked to rate themselves on

answering the questions, and to think about if they communicated ideas accurately and

explained ideas in their own words (Appendix F). After, participants received the bogus

"computer-scored" negative feedback -- a score of 2/5. As described above, the 2/5 score

was accompanied by the randomly-assigned absolute (Appendix G) or the relative

(Appendix H) description of the score. Next, also through random assignment, one group

of participants completed the reflection activity (Appendix I), while the other group

completed the time filler task (Appendix J). After, participants completed the PANAS a

second time (Appendix K) in order to determine if reflection could mitigate negative

emotion. Participants also completed The Feedback Acceptance Scale (Appendix L).

Lastly, participants completed the Response Check (Appendix M) and received the

Debriefing Form (Appendix N). In addition to the cover story, inclusions of seemingly

unrelated measures of testing a workplace manual were justified for other reasons related

to the research. Deception was used in this study because if participants knew that we

were investigating the effects of reflection, it may have interfered with the results.

Results

Participants

Participants had a mean age of 34.84 (range 18 to 67; $SD = 9.72$), 108 were

women. All participants were employed either part-time or full-time.

A power analysis indicated a sample size of approximately 210 would be needed

to detect a medium effect size. In order to account for careless responding, missing data,

and uneven sample sizes (that could alter the effect size), 300 participants, approximately 75 participants in each condition, were initially tested. After data cleaning, the number of participants was below what was adequate for sufficient power. As a result, another batch of 200 participants was collected. Every participant who took-part in the first batch of the study was precluded from participating in the second batch. All participants were over 18 so that they could give consent. The total N was 250 participants. There were 66 participants in the relative condition and time filler condition, 61 participants in the relative condition and reflection condition, 60 participants in the absolute condition and time filler condition, and 63 participants in the absolute condition and reflection condition.

Participants were not included in the analyses if they failed any question in the careless responding questions. If manipulation check questions were answered "strongly disagree," "disagree," or "neutral," the participant's data was deleted from the analyses. This was to ensure participants believed the manipulation. However, for the manipulation check item "My computer-scored rating came from an algorithm developed over the course of 5 years," the answer "neutral" did not result in a deleted item. This is because some participants explained that they did not remember exactly how many years it took to develop the algorithm. Perhaps, some participants thought of it as a knowledge check rather than a manipulation. As such, the "neutral" responders were kept in the analyses.

The open-ended questions following the workplace manual reading, as well as the reflection questions, were read before the hypotheses were tested. Cases were deleted if participants did not answer the question, i.e. if they wrote down something irrelevant to the posed question, confessed they did not believe the manipulation, or stated that the

computer-scored rating was merely an error. In addition, answers to the final question, "Is there anything you would like to say about this MTurk study?" were read to determine if anyone stated they did not believe the manipulation. All open-ended questions were read twice to confirm the cases that were deleted should have been deleted, and if any additional cases should have been deleted.

Preliminary Analyses

A mean negative affect (NA) score was calculated using all 10 negative items from the PANAS, distressed, upset, guilty, scared, hostile, irritable, ashamed, nervous, jittery, and afraid (items #2, #4, #6, #7, #8, #11, #13, #15, #18, and #20, respectively, in Appendix D and K). It was specified that at least 80 percent of the NA items had responses to create a mean for each participant. No mean NA scores were missing. There was a significant difference between the mean NA scores before the manipulation (M =1.34, SD = 0.49) and after the manipulation (M =1.40, SD = 0.61), although the actual mean difference values were small, $t(249)$ = -6.49, $p < .001$, d = -0.29. Due to the non-normal distribution, this was also confirmed using a Wilcoxon Signed Ranks test, Z = -6.87, $p < .001$.

The NA items were skewed because the task was not extremely negative, and most participants rated themselves on the lower end of the scale. However, ANOVA's are still robust despite violating this assumption.

Test of Hypotheses

A 2X2 Analysis of Variance (ANOVA) was carried out on two independent variables (reflection activity, feedback presentation method) on the level of NA. Reflection activity had two levels, the reflection and time filler condition. Feedback

presentation method also had two levels, the absolute and relative condition. The main effect of the reflection activity was investigated. This was used to determine if H_1 was rejected or confirmed: when individuals received negative feedback, did the level of NA they feel depend on if they participated in the reflection activity or not? It was also investigated if there was a main effect of the feedback presentation method or if there was an interaction effect between the conditions. This is because when evaluations use comparative language, they usually impact someone more negatively, as the feedback is directed at the 'self level' (Kluger & DeNisi, 1998).

In addition, self-evaluation scores were tested as a co-variate to determine if these ratings would affect the relationship between the independent and dependent variables. Previous research has found that a higher discrepancy between a given rating and self-rating predicts a negative reaction. ANCOVA was used to analyze this, with the self-rating used as a covariate, because the computer-scored rating of 2/5 was kept constant for all participants. This analysis was carried out to test H_2: that the self-rating was a significant covariate.

H_3 was tested using ANOVA to see if reflection would result in higher scores on the self-awareness and intent to use subscales. All subscales were used as a dependent variable in the ANOVA for exploratory purposes. This analysis would also allow us to investigate if the feedback presentation method had an effect on the feedback acceptance subscales, and if there was an interaction between the conditions.

Test of H_1

The ANOVA revealed that there was no main effect of the reflection activity, $F(1, 246) = 1.83$, $p = .18$, $\eta^2 = .007$ on the mean NA scores, indicating no significant

difference between the reflection condition ($M = 1.45$, $SD = 0.61$) and time filler

condition ($M = 1.35$, $SD = 0.53$). There was also no main effect of the feedback

presentation method, $F(1, 246), = 1.17$, $p = .28$, $\eta^2 = .005$, indicating no significant

difference between the absolute condition ($M = 1.36$, $SD = 0.56$) and the relative

condition ($M = 1.45$, $SE = 0.66$). There was also no significant interaction, $F(1, 246) = 0.19$, $p = .66$, $\eta^2 = .001$. Therefore, H_1 was rejected. The means and standard deviations

are displayed in Table 1.

Table 1.
Means and Standard Deviations of Mean NA Scores

Condition	Reflection		Time Filler	
	M	SD	M	SD
Absolute	1.40	0.56	1.33	0.54
Relative	1.51	0.77	1.38	0.53

M = Mean, SD = Standard Deviation

Follow-Up Analyses for H_1

We surmised that some of the NA items in the PANAS correspond to emotions

that could not reasonably be affected because of the nature of this study. For example,

"ashamed" seemed unlikely to apply because participants knew that their responses

would be completely confidential. Thus, as a follow-up, five graduate students rated how

"amenable to change" the NA items were given the study's tasks, that is, which emotions

would be susceptible to change pre-post the tasks in the study. The subset of NA items

chosen were based both of the graduate students' ratings as well as considering the

realities and relevance in the study context. All items had high "amenable to change"

ratings as judged by the graduate students, and were further justified due to conceptual

grounds. Upset was chosen because it was logical to assume participants would feel upset

when receiving negative feedback, as this may cause one to be unhappy. Irritable and distressed were chosen, as they may be associated with anxiousness – which could have been felt after receiving negative feedback – and agitation – which could have also resulted from the negative feedback. Hostile was chosen because participants may have felt bitter or opposed to the feedback, and this emotion may have captured that. The subset of NA items was judged purely conceptual (not empirical) grounds, and were grouped together and used as the dependent variable: distressed, upset, hostile, and irritable (#2, 4, #8 and, #11 respectively, in Appendix D and K).

The mean score of the subset of NA items before and after the manipulation was significantly different, $t(249) = -7.703$, $p < .001$, $d = -0.43$. This was confirmed with the Wilcoxon Signed Ranks test, $Z = -7.42$, $p < .001$, due to the skewed distribution.

The ANOVA revealed that there was a main effect of the reflection activity, $F(1, 246) = 5.20$, $p = .023$, $\eta^2 = .021$, indicating a significant difference between the reflection ($M = 1.63$, $SD = 0.77$) and the time filler condition ($M = 1.43$, $SD = 0.60$), on the subset of NA items. The mean differences between the two groups before and after the manipulation are displayed in Figure 1.
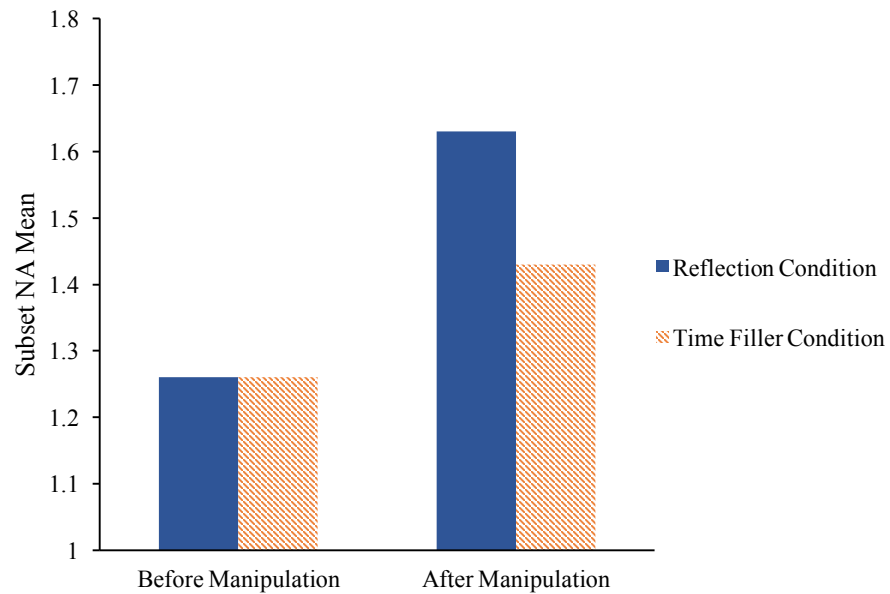
Figure 1.
Means of NA Subset in Reflection and Time Filler Conditions Before and After Manipulation

There was no main effect of the feedback presentation method, $F(1, 246) = 1.61$, $p = .21$, $\eta^2 = .006$, indicating no significant difference between the absolute ($M = 1.48$, $SD = 0.65$) and relative condition ($M = 1.58$, $SD = 0.74$). There was no interaction effect, $F(1, 246) = 0.21$, $p = .64$, $\eta^2 = .001$. Means and standard deviations are displayed in Table 2.

Table 2.
Means and Standard Deviations of Subset NA scores by Condition

Condition	Reflection		Time Filler	
	M	SD	M	SD
Absolute	1.55	0.71	1.40	0.57
Relative	1.70	0.83	1.47	0.62

M = Mean, SD = Standard Deviation

Test of H₂

ANCOVA was also used to test if the self-evaluation rating was a covariate in predicting mean NA scores, but was not significant, $F(1, 245) = 0.31, p = .58, \eta^2 = .001$. This was also tested with the subset of NA items (distressed, upset, irritable, and hostile) and it was not significant, $F(1, 245) = 2.13, p = .15, \eta^2 = .009$. Therefore, H$_2$ was rejected, meaning that participants self-evaluations did not significantly influence the effect of the manipulations.

Test of H$_3$

ANOVA was carried out with the Feedback Acceptance Subscales as the dependent variables. All means and standard deviations are displayed in Table 3. First, self-awareness was analyzed. There was no main effect of the reflection activity, $F(1, 246) = 0.02, p = .90, \eta^2 < .001$, no main effect of feedback presentation method (relative/absolute), $F(1, 246) = 0.312, p = .58, \eta^2 = .001$, and no interaction effects, $F(1, 246) = 0.21, p = .64, \eta^2 = .001$. Thus, H$_{3a}$ was rejected, as there were no higher ratings of self-awareness for participants in the reflection condition.

Intent to use was then tested. There were no main effects of the reflection activity, as well as feedback presentation method (absolute/relative), $F(1, 246) = 0.08, p = .78, \eta^2 < .001; F(1, 246) = 0.46, p = .50, \eta^2 = .002$, respectively. There was also no interaction, $F(1, 246) = 1.97, p = .16, \eta^2 = .008$. Therefore, H$_{3b}$ was also rejected, as those participants in the reflection condition did not rate intent to use higher.

For exploratory purposes, fairness and clarity were also used as dependent variables. There was no main effect of the reflection activity on fairness, $F(1, 246) = 0.99, p = .32, \eta^2 = .004$, no main effect of feedback presentation method (absolute/relative), $F(1, 246) = 0.67, p = .42, \eta^2 = .003$ and no interaction, $F(1, 246) =$

1.83, $p = .18$, $\eta^2 = .007$. Interestingly, there was a main effect of the reflection activity on

clarity, $F(1, 246) = 3.71$, $p = .055$, $\eta^2 = .015$, such that those participants in the reflection

condition rated the clarity of the feedback as lower; i.e. they found the computer-scored

rating less clear. However, this finding was marginally significant and had a small effect

size, so must be interpreted with caution. There was no main effect of feedback

presentation method (absolute/relative), $F(1, 246) = .30$, $p = .58$, $\eta^2 = .001$, nor was there

an interaction, $F(1, 246) = 1.53$, $p = .22$, $\eta^2 = .006$.

Table 3.
Means and Standard Deviations of Feedback Acceptance Subscales by Condition

	Feedback Acceptance Questionnaire Subscale	Reflection Condition		Time Filler Condition	
Absolute Condition		M	SD	M	SD
	Self-Awareness	2.88	1.08	2.81	1.02
	Fairness	2.59	1.14	2.54	1.00
	Clarity	3.04	1.15	3.14	1.11
	Intent to Use	2.86	1.19	2.70	1.17
Relative Condition					
	Self-Awareness	2.75	1.08	2.79	1.00
	Fairness	2.30	1.03	2.62	1.04
	Clarity	2.79	1.15	3.23	1.01
	Intent to Use	2.55	1.21	2.80	1.10

M = Mean, SD = Standard Deviation

As seen in Table 4, the means of these items indicated that most participants did

not find that the feedback promoted self-awareness, i.e. that it allowed them to learn more

about themselves; that the feedback was fair or clear; and that the feedback would be

thought about or useful for the future. This may have occurred given that there was little

qualitative feedback presented to participants. Notably, the fairness subscale had the

lowest mean in each condition, and had the lowest overall mean, ($M = 2.51$, $SD = 1.06$).

This indicated that many participants either disagreed or felt neutral when answering the fairness items.

Table 4.
Overall Means and Standard Deviations of Feedback Acceptance Subscales

Feedback Acceptance Questionnaire Subscales	M	SD
Self-Awareness	2.81	1.04
Fairness	2.52	1.05
Clarity	3.06	1.11
Intent to Use	2.73	1.17

M = Mean, SD = Standard Deviation

Discussion

Unexpectedly, reflection did not result in lower negative affect (NA) after negative feedback. There was no significant difference between the mean NA in the reflection and time filler conditions. Moreover, when a subset of NA items more likely to be affected by our manipulation was analyzed, reflection resulted in *higher scores* of NA.

One reason for these results could be that providing the numeric rating with limited qualitative feedback stifled the effects of reflection. When asked, "Is there anything else you would like to say about this MTurk study?" a few of the participants stated it was difficult to understand why they got the rating they did when there was a lack of feedback comments were provided. Perhaps reflection without detailed qualitative feedback is ineffective, and may actually produce a negative reaction. It is possible that the benefits of reflection are only evident when thorough comments and suggestions accompany feedback. Previous studies investigating feedback and reflection (Anseel et al., 2009; Phielix et al., 2011) provided more detailed qualitative feedback, which may have assisted in the effectiveness of the reflection.

Perceived justice may have been a confounding variable, as numeric feedback was provided by a "computer-scored algorithm" instead of a "real person," and feedback lacked a qualitative detail. The mean of the fairness subscale, a construct similar to justice, was such that most participants rated between disagree and neutral, suggesting that participants felt a low level of justice. Previous literature has found that justice correlates with feedback acceptance and attitudinal response (Leung, Su & Morris, 2001), demonstrating that justice could have interfered with the results.

Previous qualitative research may support this study's overall findings. One study by Hobbs (2007) examined the effects of journaling in short teaching courses. Students were asked to write in a journal and were required to include information outlined by the course. Researchers found that participants could have a negative response to reflection due to feeling that they had to please the individual reading the reflection. Another study found that people do not know the why reflection is meaningful (Roberts, 1998), which could also result in negative reactions.

The reflection activity in this study may have induced a similar effect to that of rumination. Self-focused rumination is defined as "…thoughts and behaviors that focus the individual's attention on the negative mood, the causes and consequences of this mood, and self-evaluations related to the mood," (Rusting & Nolen Hoeksema, 1998, p. 790). Rusting and Nolen Hoeksema (1998) found that rumination increased feelings of anger when compared to a distractor group. In the rumination group, participants were asked to respond to either self-focused items or emotion-focused items, that is, items pertaining to anger and related thoughts, while the distractor group was asked to think about something completely neutral and irrelevant to the study. This study built on

previous research that found rumination had increased sadness and anxiety (Rusting &

Nolen Hoeksema, 1998). The reason rumination increased these negative emotions was

justified through spreading activation theory: when an emotion is felt, previous thoughts

linked to that feeling are also recalled and experienced as they are all intertwined in a

neural network, which causes the emotion to last longer. The same effect may have

occurred in our reflection condition: the reflection may have activated a neural network,

resulting in a higher score on pertinent NA items.

Goal setting may thus be the most effective way for employees to accept negative

feedback. By setting goals, employees can focus on developing specific skills (Locke &

Latham, 2002), which may increase feedback acceptance by refocusing attention from

criticism to improvement of future performance. In Phielix et al.'s study (2011),

reflection incorporated aspects of goal setting, a possible explanation as to why it was

effective.

There was no significant difference between the absolute and relative conditions

on NA. This may have occurred because the comparison was not salient; participants may

have been unaware of the abilities and identities of the individuals they were being

compared to. The results may have been different if this were carried out in a workplace

setting, where people have a relationship with their coworkers and know their abilities. In

fact, according to the Self-Evaluation Maintenance (SEM) model of social behaviour,

people feel inclined to improve or uphold their current self-evaluation, which is strongly

affected by: a) performance feedback on a task that is relevant (versus irrelevant) to one's

self-definition; and b) when ability on a task is compared to people "close" to that

individual – friends, family members, co-workers – rather than to those that are distant –

strangers (Tesser, Millar & Moore, 1987). Research by Salovey and Rodin (1984) investigated this phenomenon, and discovered that anxiety and depression was higher among those participants who received negative feedback on relevant tasks, and when feedback was compared to close others. This is because negative emotions result from one's self-evaluation being threatened. These findings demonstrate that participants may have been less affected by the relative condition than predicted because the comparison was made to "distant" others.

The second hypothesis, that the self-evaluation rating would be a significant co-variate in the study was rejected. Participants may have been less defiant about their self-rating because it was based on one circumscribed task and not on overall job performance. Thus, they were not afforded time to develop a sense of their abilities. This relates to the SEM model in that participants may have been less affected by the feedback, as the task was not pertinent to participants' self-definition. In addition, the feedback provided in this study came from a "computer-scored algorithm," which may have affected how negatively participants felt from the score. They may have felt that the score was less "real," and justified that the result was from a computer which cannot analyze their actual performance. Therefore, the score may have had less of an effect on participants regardless of their self-rating; in turn, the self-rating would not have been a significant co-variate.

In the exploratory analyses, one subscale of the Feedback Acceptance Scale, clarity, had a marginally significant *negative* main effect of the reflection activity. Participants in the reflection condition were required to write about their performance and computer-scored feedback, explaining why they may have received the rating they did.

Perhaps, participants then realized the feedback lacked clarity, as there was a numeric score but little qualitative feedback to further guide the reflection. As a result, these participants may have rated lower on clarity.

Self-awareness and intent to use did not significantly differ based on condition. This may have occurred for the same reason clarity was significant: because feedback was numeric with little qualitative content.

Limitations & Future Research

Previous research (Phielix et al., 2011; Anseel et al., 2009) contained varying numbers of reflection items in the reflection activity. In future, researchers should examine how many items to include in a facilitated reflection, as well as the optimal duration for a reflection activity.

Only four subscales from the Feedback Acceptance Scale were used because they were the only scales relevant to this study. Other studies could use the entire Feedback Acceptance Scale to investigate how reflection influences all subscales.

A limitation to this study is generalizability. This study took place over a relatively short period of time and participants were aware that it was not a real job. This study was a lab study, and the feedback and reflection activity were based on a question and answer exercise, not overall work performance. This was necessary in order to gain the kind of control that the experimental method allows. However, participants may have been less affected by the feedback, as well as the reflection, since it was based on a task they may have not felt was as important as much as a job performance. Because this study was not carried out in the workplace, it also made the absolute and relative conditions' null results not as generalizable. In the workplace, employees are aware of

who they are being compared to, and have ideas about both their own and others' abilities. According to the SEM model, having the tasks be more relevant to one's self-definition and being compared to people "closer" to oneself, may yield different results. Future research in field settings would also be advisable.

In future, it would be interesting to test the impact of motivation on the effectiveness of reflection. To my knowledge, this has not been examined in the workplace setting. Additionally, research should be conducted to determine if reflection is effective in a coaching setting. This is a setting in which both the employee and employer are focused on performance improvement. As such, reflection may be helpful in learning.

As mentioned, previous research (Hobbs, 2007) found that when reflection activities were forced, participants felt their responses should please the reader, and that participants did not have a favourable opinion of reflection, as it appeared to have little value (Roberts, 1998). This could be addressed in future research. Perhaps after a reflection activity, participants can be asked if the activity was helpful or not. The groups could then be compared. Hobbs (2007) also suggested gradually introducing reflection activities to prevent feelings of resentment. This could also be investigated in future.

Lastly, future research could examine both the role of emotion and learning to get a full understanding of how each lead to performance changes after reflection in the workplace. Perhaps emotion does not manifest as expected, as reflection may not result in immediate decreases in negative emotion, yet still have benefits to learning. Alternatively, reflection could simply increase negative emotions and have detrimental effects on being receptive to negative feedback. Therefore, it is imperative to do follow-

up studies on emotion, feedback, and reflection to understand how emotion manifests itself, and make a theoretical contribution about both the effects of reflection and the benefits of the feedback process.

Conclusion

Although there are theoretical claims that reflection has the potential to result in the ability to manage emotions, understand new perspectives, and result in performance change, we found evidence that reflection may not be all it is cracked up to be in the case of performance feedback. The reflection activity did not result in less negative emotion; in fact, there was evidence it actually increased negative emotion. Factors such as how feedback was delivered (without little qualitative content), worries about pleasing the researcher, the belief that reflection was not helpful, and experiencing effects similar to that of rumination, may have caused participants to feel slightly more negative after completing the reflection activity. Future research on reflection in different contexts, as well as the effect reflection has on both emotion and learning from feedback should be examined. This would contribute to the theoretical underpinnings of reflection and further uncover the usefulness of reflection in a work setting.

References

Anseel, F., Lievens, F., & Schollaert, E. (2009). Reflection as a strategy to enhance task performance after feedback. *Organizational Behavior and Human Decision Processes*, *110*(1), 23-35.

Ashford, S. J. (1986). Feedback-seeking in individual adaptation: A resource perspective. *Academy of Management Journal, 29*(3), 465-487. doi:http://dx.doi.org/10.2307/256219

Belschak, F. D., & Den Hartog, D. N. (2009). Consequences of positive and negative feedback: The impact on emotions and extra-role behaviors. *Applied Psychology*, *58*(2), 274-303.

Boud, D., Keogh, R., & Walker, D. (1985). Promoting reflection in learning: A model. *Reflection: Turning Experience into Learning*, 18-40.

Boud, D. (2001). Using journal writing to enhance reflective practice. *New directions for adult and continuing education*, *2001*(90), 9-18.

Brett, J. F., & Atwater, L. E. (2001). 360° feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology, 86*(5), 930-942. doi:10.1037/0021-9010.86.5.930

DeRue, D. S., Nahrgang, J. D., Hollenbeck, J. R., & Workman, K. (2012). A quasi experimental study of after-event reviews and leadership development. *Journal of Applied Psychology*, *97*(5), 997.

Dipboye, R. L., & De Pontbriand, R. (1981). Correlates of employee reactions to performance appraisals and appraisal systems. *Journal of Applied psychology*, *66*(2), 248.

Hobbs, V. (2007). Faking it or hating it: can reflective practice be forced?. *Reflective practice*, *8*(3), 405-417.

Kedharnath, U., Garrison, L., & Gibbons, A. M. (2010). A multidimensional measure of feedback acceptance. Presented at the 25th Annual Meeting of the Society for Industrial Organizational Psychology in Atlanta, GA.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin*, *119*(2), 254.

Kluger, A. N., & DeNisi, A. (1998). Feedback interventions: Toward the understanding of a double-edged sword. *Current directions in psychological science*, *7*(3), 67-72.

Kolb, D. A., Boyatzis, R. E., & Mainemelis, C. (2001). Experiential learning theory: Previous research and new directions. *Perspectives on thinking, learning, and cognitive styles*, *1*, 227-247.

Lee, K., & Ashton, M. C. (2016). Psychometric properties of the HEXACO-100. *Assessment*, 1073191116659134.

Leung, K., Su, S., & Morris, M. W. (2001). When is criticism not constructive? The roles of fairness perceptions and dispositional attributions in employee acceptance of critical supervisory feedback. *Human Relations*, *54*(9), 1155-1187.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist*, *57*(9), 705.

Loo, R., & Thorpe, K. (2002). Using reflective learning journals to improve individual

and team performance. *Team Performance Management: An International Journal*, *8*(5/6), 134-139.

Phielix, C., Prins, F. J., & Kirschner, P. A. (2010). Awareness of group performance in a CSCL-environment: Effects of peer feedback and reflection. *Computers in Human Behavior*, *26*(2), 151-161.

Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., & Jaspers, J. (2011). Group awareness of social and cognitive performance in a CSCL environment: Effects of a peer feedback and reflection tool. *Computers in Human Behavior*, *27*(3), 1087-1102.

Roberts, J. (1998). *Language teacher education*. Routledge.

Rusting, C. L., & Nolen-Hoeksema, S. (1998). Regulating responses to anger: effects of rumination and distraction on angry mood. *Journal of personality and social psychology*, *74*(3), 790.

Salovey, P., & Rodin, J. (1984). Some antecedents and consequences of social-comparison jealousy. *Journal of Personality and Social Psychology*, *47*(4), 780.

Sargeant, J., Mann, K., Sinclair, D., Van der Vleuten, C., & Metsemakers, J. (2008). Understanding the influence of emotions and reflection upon multi-source feedback acceptance and use. *Advances in Health Sciences Education*, *13*(3), 275-288.

Scanlan, J. M., & Chernomas, W. M. (1997). Developing the reflective teacher. *Journal of advanced nursing*, *25*(6), 1138-1143.

Tesser, A., Millar, M., & Moore, J. (1988). Some affective consequences of social

comparison and reflection processes: the pain and pleasure of being close. *Journal of personality and social psychology*, *54*(1), 49.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, *54*(6), 1063.

World Diamond Council in Alliance with Jewelers Vigilance Committee, Jewellers of America, Diamond Dealers Club, Diamond Manufacturers and Importers Association of America, and Jewelry Information Center. *Employee Training Manual: How to Train Your Employees about Conflict Diamonds.* Retrieved from http://www.diamondfacts.org/pdfs/industry/TrainingGuide_v5.pdf

Xie, Y., Ke, F., & Sharma, P. (2008). The effect of peer feedback for blogging on college students' reflective learning processes. *The Internet and Higher Education*, *11*(1), 18-25.

Appendix A

Cover Story

Pay: $2 We are looking for employed participants. We are testing the effectiveness of a workplace manual for a retail company that sells diamonds. This workplace manual is a training manual to prepare employees for customers' questions regarding conflict diamonds. Many measures included are to test the different aspects of the training manual and the activities associated with it. In addition, you will be asked to fill out several measures that we need to include as the responses may affect the results. The measures may also be included for further research.

Appendix B

Letter of Information and Consent Form

Project Title: Investigating the Effectiveness of a Workplace Manual
Principal Investigator: *Richard Goffin, Professor, The University of Western Ontario*

1. **Invitation to Participate**

   You are being invited to participate in this research study investigating the effectiveness of a workplace manual and activities associated with the workplace manual.

2. **Purpose of the Letter**

   The purpose of this letter is to provide you with information required to make an informed decision regarding participation in this research.

3. **Purpose of this Study**

   The purpose of this study is to allow researchers to learn more about the effectiveness of a workplace manual and activities associated with the workplace manual.

4. **Inclusion Criteria**

   Individuals who are English speaking are eligible to participate in this study. Both females and males can participate in this study. Individuals can only participate in this study if they are currently employed.

5. **Exclusion Criteria**

   Individuals who are unable to see or hear are not eligible to participate in this study. This is because the study involves an online questionnaire where participants will have to read and answer questions. Individuals must also be over the age of 18 so they can consent to participate.

6. **Study Procedures**

   If you agree to participate, you will first be asked to fill out a short participant information form, as well as a measure that asks you about your feelings. Then you will be asked to read a workplace training manual about conflict diamonds. You will be asked to answer open-ended questions on the content in the manual. You will then evaluate yourself on how well you communicated ideas when answering the questions. Your answers will also be scored using a computer-based machine learning algorithm which will provide you with feedback. After,

you will be asked to complete additional questionnaires to assess the overall effectiveness of the training manual, as well as complete measures that provide researchers with additional information about the manual. You will also receive the debriefing form. The whole study should not take longer than 30 minutes. The entire study will be online, and will take only one session to complete. There will be a total of 500 participants in this study.

7. **Possible Risks and Harms**

You may experience negative affect. There are no other known risks or discomforts associated with participating in this study.

8. **Possible Benefits**

The possible benefits to society may be that this study will help practitioners decide what type of manual and training exercises are the most effective. It could also contribute to literature on workplace feedback since the training exercises include self-evaluations and feedback.

9. **Compensation**

You will be compensated $2.00 for your participation.

10. **Voluntary Participation**

Participation in this study is voluntary. You may refuse to participate, refuse to answer any questions or withdraw from the study at any time. You do not waive any legal rights by consenting to this research.

11. **Confidentiality**

All data collected will be anonymized. All electronic data will be kept on a password protected, encrypted file. To allow for independent verification of results, all data will be kept for a minimum of 5 years after the publication of the final article that is based on this research. The completely anonymized data may eventually be posted on an open access site for other researchers' access. Representatives of The University of Western Ontario Non-Medical Research Ethics Board may contact you to monitor the conduct of the research.

12. **Contacts for Further Information**

If you require further information regarding this research project or your participation in this study you may contact the supervisor, Richard Goffin, at ▆▆▆▆▆▆▆▆ for principle investigator or the student researcher, Rebecca Factor, at ▆▆▆▆▆▆▆. If you have any questions about your rights as a research participant or the conduct of this study, you may contact The Office of Research Ethics ▆▆▆▆▆▆▆, email: ▆▆▆▆▆▆.

13. Publication

If the results of the study are published, your name will not be mentioned.

<u>By clicking 'Next,' you are giving consent to participate in this study.</u>

Appendix C

Participant Information Form

1. Gender
   o Male
   o Female
   o Other

2. Age: _____

3. First Language: _____

4. I am employed:
   o Part time, with one job (24 hours or less at one job)
   o Part time, with two or more jobs (24 hours or less at each job)
   o Full time (25 hours or more at one job)
   o Not employed
   o Other (Please explain) _____

5. I have completed (click all that apply):

   High School
   o High school diploma

   College
   o Please specify highest level of completion (ex: <u>completed</u> first year): _____

   o College Graduate

   University
   o Please specify highest year of completion (ex: <u>completed</u> first year): _____

   o University Graduate, please specify (Bachelor, Master's, PHD, other): _____

   Other
   o Please Specify: _____

6. What type of work do you do, or have you done?
   o Management
   o Office/Administrative
   o Business and Financial Operations
   o Architecture/Engineering
   o Art and Design
   o Entertainer/Performer

- o Media and Communications
- o Computer/Mathematical
- o Farming/Fishing/Forestry
- o Building, Grounds, Cleaning, and Maintenance
- o Life Science
- o Physical Science
- o Military or Protective Service
- o Healthcare Practitioners or Technician
- o Healthcare Support
- o Community and Social Service
- o Social Science
- o Legal Occupations
- o Education/Training/Library
- o Transportation
- o Personal Care and Service
- o Construction/Installation/Repair
- o Food Preparation/Serving
- o Sales
- o Production/Manufacturing

7. Have you ever had to evaluate an employee's job performance and/or give job performance feedback?
- o No
- o Yes

8. Have you ever had a supervisory role?
- o No
- o Yes
  - ➢ If yes, for how long?
    - o Less than 2 months
    - o 2 months to 6 months
    - o 6 months to 1 year
    - o 1 year to 5 years
    - o 5 years or more

9. Have you ever been given job feedback and/or had your job performance rated?
- o No
- o Yes

Appendix D

Feelings Scale
(Time 1)

We find that people's feelings may affect their responses in this study, so we are
including the following questions to find out more about your feelings. This scale will be
included twice in this study.

This scale consists of a number of words that describe different feelings and emotions.
Please read each word and then mark the appropriate answer next to that word. Indicate
to what extent *you feel this way right now, that is, at the present moment*.

Very slightly or not at all	A little	Moderately	Quite a bit	Extremely

Interested _____
Distressed_____
Excited _____
Upset _____
Strong _____
Guilty _____
Scared _____
Hostile _____
Enthusiastic _____
Proud _____
Irritable _____
Alert _____
Ashamed _____
Inspired _____
Nervous _____
Determined _____
Attentive _____
Jittery _____
Active _____
Afraid _____

Appendix E

Conflict Diamonds Training Manual

For this task, please imagine you are an employee at an American retail store that sells diamonds. The manager has scheduled a meeting to discuss the issue of "conflict diamonds," that is, diamonds that are mined where there is armed conflict and the diamonds are traded to fund the attacks and violence. Below is the training manual you receive to deal with customers' concerns regarding conflict diamonds. Please read the training manual carefully. After you have had a chance to read it, you will be asked some questions about it.

Consumers are more socially aware so we need to let them know we do not support or sell conflict diamonds. They also need to know that for years the industry has been actively working to eliminate conflict diamonds from the marketplace.

I want to assure you that every diamond we sell is warranted by our suppliers to be from sources that are free from conflict. We are part of an industry-wide, global effort to eradicate illegal diamonds. We will be discussing how you can speak confidently to your customers and assure them that the diamonds they are buying come from legitimate sources. Diamonds are a lifeline and provide a livelihood for many people in Africa and across the globe, so we also need to do our part to ensure that our customers understand that diamonds are vital to the stable societies of Africa and to the growth and future of many nations.

In 1998, our industry became aware that rebel movements in Africa were selling, among other things, illegally obtained rough diamonds to pay for their wars. As awareness grew, it became apparent that this illegal trade and the rebel wars it supported were causing a lot of human suffering. The diamond industry realized that this trade had to be stopped.

It took four years to create, ratify and adopt a process that prohibits conflict diamonds from entering the legitimate rough diamond supply. On November 5, 2002, there were 52 governments who ratified the process. This process is called the Kimberley Process.

The Kimberley Process sets rules for the import and export of rough diamonds. Every shipment of diamonds that crosses an international border must be certified, numbered and sealed. It is a system that includes tamper-resistant containers and forgery-resistant certificates. The United States Government has a law that enforces the Kimberley Process, called the Clean Diamond Trade Act. It requires annual reviews of the standards, practices and procedures of the systems in place in the United States, which imports, trades or exports rough diamonds.

The Kimberley Process is helping put an end to the problem of conflict diamonds. When the problem was at its height, conflict diamonds accounted for no more than 4% of the world's diamond supply. Now, they represent considerably less than 1%. And the industry is committed to totally eliminating them. The Kimberley Process guarantees that only diamonds from legitimate sources are used in our jewelry. While we may not know exactly where a diamond originated, you can guarantee that every diamond we sell was handled within the Kimberley Process and therefore is certified to be from sources that are free from conflict.

The diamond industry adopted a System of Warranties to further assure consumers of the origin of their diamonds. This system requires every buyer and seller of polished diamonds and diamond jewelry to make the following statement on all invoices:

"The diamonds herein invoiced have been purchased from legitimate sources not involved in funding conflict and in compliance with United Nations resolutions. The seller hereby guarantees that these diamonds are conflict free, based on personal knowledge and/or written guarantees provided by the supplier of these diamonds."

—The System of Warranties Assurance Statement

In addition to putting this statement on every invoice, businesses are required to keep records of their invoices and to have them audited. Government agencies can also request proof of compliance.

We are in full compliance with this process, and we require every one of our vendors to provide us with written assurances that every diamond we buy has been sourced legitimately.

We are totally committed to being able to assure our customers that their diamonds are certified to be from sources that are free from conflict.

When a customer says, "How can I trust that these diamonds are conflict free?" you will say...

"I can assure you that all diamonds here at this jewelry store are from areas that are free of conflict—because we only source our diamonds from suppliers that are in compliance with the Kimberly Process and we participate in the System of Warranties. Plus it is illegal in the U.S. to sell conflict diamonds—and we follow the law."

"Did you know that the U.S. is one of over 68 countries that have laws to ensure that no conflict diamonds cross our borders?"

"Before these laws were enacted, the amount of conflict diamonds was estimated to be 4% of the world's diamonds; today that number is considerably less than 1%. Our industry will not rest until it is 0%."

"While there may be people who are engaged in illegal diamond trading, it is important to deal with a jeweler you know and trust. We have been in business for decades. We have also been part of this community for decades. We require official statements of assurance from every single one of our vendors."

"We have written assurances from every vendor that our diamonds are legally and legitimately sourced."

In the following questions, please answer to the best of your ability and in **your own words**. We do NOT want you to use the exact same wording that was in the training manual.

Questions:

1. If a customer tells you they are concerned that the diamonds for sale in your store are conflict diamonds, how would you respond? Please provide your complete response.

2. If a customer asked you for your opinion on conflict diamonds, how would you respond? Please provide your complete response.

3. What would you do if you notice that a customer does not want to buy jewelry that contains diamonds, and you think that this is because they have concerns about conflict diamonds? What would you say to them? Please provide your complete response.

Appendix F

Self-Evaluation of Your Answers to the Questions

We would like to get an idea of how well you think you answered the questions. In many workplaces, employees have the opportunity to complete a self-evaluation as well as receiving other feedback on their performance, so we would like to follow a similar procedure here.

Please rate yourself from 1-5 on how well you responded to the questions. Think about if you communicated ideas accurately and explained ideas in your own words.

Use this scale when deciding your rating:

1	2	3	4	5
Poor		**Satisfactory**		**Excellent**

A rating of **1** (out of 5) would indicate that your performance is extremely poor.

A rating of **2** (out of 5) would indicate that your performance is better than poor, but less than satisfactory.

A rating of **3** (out of 5) would indicate that your performance is satisfactory.

A rating of **4** (out of 5) would indicate that your performance is better than satisfactory, but less than excellent.

A rating of **5** (out of 5) would indicate that your performance is absolutely excellent.

Please rate your own performance on answering the questions: _____

Appendix G

Computer-Scored Evaluation of Your Answers to the Questions
(Absolute Condition)

While you were completing your self-rating, your answers were being computer-scored using a machine learning algorithm developed over the course of 5 years. This was used to objectively determine your true performance in responding to the questions. We wanted to compare this to your self-rating to help us get a fuller picture of your performance.

Your Computer-Scored rating is:

** 2 out of 5 **

As a reminder, this was the scale used:

1	2	3	4	5
Poor		**Satisfactory**		**Excellent**

This means that your performance, **2 out of 5,** was **better than poor, but less than satisfactory**.

Appendix H

Computer-Scored Evaluation of Your Answers to the Questions
(Relative Condition)

While you were completing your self-rating, your answers were being computer-scored using a machine learning algorithm developed over the course of 5 years. This was used to objectively determine your true performance in responding to the questions. We wanted to compare this to your self-rating to help us get a fuller picture of your performance.

Your Computer-Scored rating is:

2 out of 5

As a reminder, this was the scale used:

1	2	3	4	5
Poor		**Satisfactory**		**Excellent**

Of all participants with a similar background to you, **64% of participants received a higher score than you did.**

Appendix I

Reflection Activity
(Follow-Up Activity)

The following are questions about the computer-scored ratings you received on your answers to the questions. Please respond honestly.

(1) The rating you gave yourself was (insert rating). Please explain your reasons for giving yourself this rating. Please try to provide at least 2 sentences or about 20 words in your response.

(2) **First**, think about the rating you gave yourself, and the computer-scored rating that you received.
Second, notice whether there was a difference in those two ratings. Also, notice how large that difference was.
Third, do you agree or disagree with the computer-scored rating you received? Why or why not?
Please try to provide at least 2 sentences or about 20 words in your response.

Appendix J

Time Filler Task

We have found that a person's likes, dislikes, and inclinations, may affect this research.

Below, you will find a series of statements that will help us learn more about your likes, dislikes, and inclinations. Please read each statement and decide how much you agree or disagree with that statement. Then, click the best response next to the statement.

strongly agree   agree   neutral (neither agree nor disagree)   disagree   strongly disagree

Please answer every statement, even if you are not completely sure of your response.

1. I would be quite bored by a visit to an art gallery.
2. I would like a job that requires following a routine rather than being creative.
3. I wouldn't spend my time reading a book of poetry.
4. I enjoy looking at maps of different places.
5. I would enjoy creating a work of art, such as a novel, a song, or a painting.
6. If I had the opportunity, I would like to attend a classical music concert.
7. I would be very bored by a book about the history of science and technology.
8. When I'm in a group of people, I'm often the one who speaks on behalf of the group.
9. People have often told me that I have a good imagination.
10. I prefer jobs that involve active social interaction to those that involve working alone.
11. Sometimes I like to just watch the wind as it blows through the trees.
12. I've never really enjoyed looking through an encyclopedia.
13. I don't think of myself as the artistic or creative type.
14. I find it boring to discuss philosophy.
15. I prefer to do whatever comes to mind, rather than stick to a plan

Appendix K

Feelings Scale
(Time 2)

We find that people's feelings may affect their responses in this study, so we are including the following questions to find out more about your feelings. We are asking you to fill out this questionnaire a second time just in case any of your feelings have changed. Please be honest.

This scale consists of a number of words that describe different feelings and emotions. Please read each word and then mark the appropriate answer next to that word. Indicate to what extent *you feel this way right now, that is, at the present moment*.

Very slightly   A little   Moderately   Quite a bit   Extremely
or not at all

Interested _____
Distressed_____
Excited _____
Upset _____
Strong _____
Guilty _____
Scared _____
Hostile _____
Enthusiastic _____
Proud _____
Irritable _____
Alert _____
Ashamed _____
Inspired _____
Nervous _____
Determined _____
Attentive _____
Jittery _____
Active _____
Afraid _____

Appendix L

Feedback Questionnaire

Now, we'd like you to respond to the following questions. We will take your responses into account when deciding whether we should include the computer-scored rating in our training program. Please answer honestly.

NOTE: IN THE QUESTIONS THAT FOLLOW THE COMPUTER-SCORED RATING IS REFERRED TO AS FEEDBACK.

Answer each item by choosing the most accurate option.

strongly disagree   disagree   neutral   agree   strongly agree

1. The feedback I received taught me something about myself.

2. I believe I am now more aware of my skill strengths.

3. I believe I will be more aware of my performance on these skills in the future.

4. I believe that the feedback criteria are fair.

5. I feel that this feedback process has been fair.

6. The procedures used to evaluate my performance were fair.

7. The feedback is easy to understand.

8. The way the feedback is presented makes sense.

9. The feedback was well organized.

10. I am likely to consider this feedback the next time I am working on tasks using a similar skill set.

11. I am likely to consider this feedback when I encounter opportunities to develop.

12. The feedback I received will influence my attitudes or effort in the future.

13. I plan on giving thought to the feedback I received in the future.

Appendix M

Response Check

In the next 4 questions, we are just checking that you are responding carefully.

1) Please choose N below
　　Y or N

2) Please choose 'agree' below
　　strongly disagree　　disagree　　neutral　　agree　　strongly agree

3) Please choose 'strongly disagree' below
　　strongly disagree　　disagree　　neutral　　agree　　strongly agree

4) The answers in this study are very important for our analyses. Do you believe we should use your data from this study for our analyses?
　　Y or N

5) The Kimberly Process:
a. is how diamonds are mined
b. is a set of rules and regulations for importing and exporting diamonds
c. none of the above

6) Conflict diamonds are diamonds:
a. that pay for wars
b. that are stolen from household break-ins
c. that are stolen from jewellers

7) The conflict diamonds training manual:
a. prepares employees on how to deal with break-ins
b. explains how to spot a conflict diamond
c. explains how to deal with customer concerns regarding conflict diamonds

Please answer the following questions honestly.

8) I was thinking about my role in helping with the training manual regarding conflict diamonds as I was responding to the questions in this study.
　　strongly disagree　　disagree　　neutral　　agree　　strongly agree

9) My computer-scored rating came from an algorithm developed over the course of 5 years.
　　strongly disagree　　disagree　　neutral　　agree　　strongly agree

10) Is there anything else you would like to say about this MTurk project?

Appendix N

Debriefing Form

Title of Research: Investigating the Effectiveness of Workplace Manuals
Investigators: Richard Goffin and Rebecca Factor

Thank you for helping us with this project--your time is much appreciated. We originally informed you that the purpose of this study was to assess the effectiveness of a workplace manual and training program. This was not the case and we apologize for deceiving you. We deceived you because, if you had known what we were studying, it could have affected the results.

The purpose of this study was to investigate the effects of reflection after receiving negative feedback. The reason this is being studied is in order to see if reflection can mitigate negative emotions and promote feedback acceptance in the workplace. This is because reflection can lead to learning, promote the understanding of a new perspective, and decrease emotional reactions. In turn, an employee would have an improved understanding and acceptance of feedback (Anseel, Lievens & Schollaert, 2009). Some participants were merely asked to complete filler items (personality questions) while other participants were asked to complete a reflection exercise after reading the manual. The computer-scored rating was NOT computer based, and everyone received the same score of 2/5. This was the case in order to examine the effects of the negative feedback.

With the results of this study, researchers may have a better understanding of what ways to improve the feedback process given that, currently, feedback often has little or no effect on future performance (Anseel et al., 2009). Individuals receiving feedback do not tend to think about the feedback very much after receiving it. Improving the feedback process could have benefits for employee development at work as employees may think about the feedback and make improvements. In turn, this could also contribute to an increase in productivity at work.

Your tasks in this study included filling out an affect questionnaire and feedback acceptance questionnaire after completing the workplace manual and receiving feedback. This gave us insight into how the negative feedback was received by participants.

If you have any further questions about this research please contact Rebecca Factor ▮▮▮▮▮▮▮▮ or the Supervisor, Dr. Richard Goffin ▮▮▮▮▮▮▮▮

If you have questions about your rights as a research participant, you should contact the Director of the Office of Research Ethics at ▮▮▮▮▮▮▮ or ▮▮▮▮▮▮▮

If you are still experiencing a negative reaction and need help, please go to http://www.yourlifecounts.org/need-help/crisis-lines to find a local helpline.

Please do NOT tell other people or post information online about the deceptions that were used in this study. This is important research and the results of this study could be affected if people are aware of these deceptions before participating.

If you would like to read more about feedback at work, please consider the sources listed below:

Anseel, F., Lievens, F., & Schollaert, E. (2009). Reflection as a strategy to enhance task performance after feedback. *Organizational Behavior and Human Decision Processes*, *110*(1), 23-35.

Dipboye, R. L., & De Pontbriand, R. (1981). Correlates of employee reactions to performance appraisals and appraisal systems. *Journal of Applied psychology*, *66*(2), 248

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin*, *119*(2), 254.

The Conflict Diamonds Training Manual was taken and adapted from (with their permission):

WORLD DIAMOND COUNCIL IN ALLIANCE WITH THE FOLLOWING U.S. ORGANIZATIONS:JEWELERS VIGILANCE

COMMITTEE, JEWELERS OF AMERICA, AMERICAN GEM SOCIETY, MANUFACTURING JEWELERS & SUPPLIERS OF AMERICA, DIAMOND DEALERS CLUB, DIAMOND MANUFACTURERS & IMPORTERS ASSOCIATION OF AMERICA, JEWELRY INFORMATION CENTER

Appendix O

Ethics Approval

Western Research

Western University Non-Medical Research Ethics Board
NMREB Delegated Initial Approval Notice

Principal Investigator: Prof. Richard Goffin
Department & Institution: Social Science\Psychology, Western University

NMREB File Number: 109450
Study Title: Investigating the Effectiveness of Workplace Manuals

NMREB Initial Approval Date: July 19, 2017
NMREB Expiry Date: July 19, 2018

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Western University Protocol		2017/07/06
Other	Cover Story / Advertisement	2017/06/12
Instruments	Debriefing Form	2017/07/06
Instruments	Follow-up Activity	2017/07/06
Letter of Information & Consent		2017/07/06
Instruments	Neutral Items	2017/06/12
Instruments	Response Check	2017/06/12
Instruments	Feedback Acceptance Questionnaire	2017/06/12
Instruments	Feelings Scale (PANAS) - Time 2	2017/06/12
Instruments	Computer-Scored Evaluation (relative)	2017/06/12
Instruments	Computer-Scored Evaluation (Absolute)	2017/06/12
Instruments	Self-Evaluation of Your Answers to Questions	2017/06/12
Instruments	Conflict Diamonds Training Manual	2017/06/12
Instruments	Feelings Scale (PANAS)	2017/06/12
Data Collection Form/Case Report Form	Demographic Questionnaire	2017/06/12

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the above named study, as of the NMREB Initial Approval Date noted above.

NMREB approval for this study remains valid until the NMREB Expiry Date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario.

Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration ▓▓▓▓▓▓▓▓▓

▓▓▓▓▓▓▓▓▓▓
Ethics Officer, on behalf of Dr. Randal Graham, NMREB Chair or delegated board member

EO: Erika Basile___ Grace Kelly___ Katelyn Harris___ Nicola Morphet___ Karen Gopaul___ Patricia Sargeant___

Curriculum Vitae

Name:	Rebecca J. Factor
Post-secondary Education and Degrees:	University of Western Ontario London, Ontario, Canada 2011-2015 B.A.
Honours and Awards:	Dean's Honor List 2014-2015
Related Work Experience:	Teaching Assistant The University of Western Ontario 2015-2017 Research Assistant The University of Western Ontario 2015

Conference Presentations:

Factor, R.J., & Goffin, R.D. (2017). *Rating Job Performance*. Presented at the annual meeting of the Canadian Psychological Association, Toronto, ON, Canada.

Daljeet, K.N., Factor, R.J., Veselka, L., & Giammarco, E. A (2017). *Predicting CWB using the Vices and Virtues Assessment*. Presented at the annual meeting of the Canadian Psychology Association, Toronto, ON, Canada.