

Evaluating the Effects of Non-Anonymity on Student Team-Member Evaluations

Mr. Taylor Robert Smith, Brigham Young University

Taylor Smith has a Bachelor of Science and a Master's degree from the civil and environmental engineering program at Brigham Young University. For his graduate studies Taylor's classwork and technical emphasis were in geotechnical engineering; however, his Master's Thesis was non-technical and he examined and tested ways to improve performance through the use of peer feedback. More particularly, he evaluated the effect that having students conduct team member evaluations non-anonymously had upon personal and team effectiveness. His research is an original work (initiated by himself), which was funded by the BYU college of engineering. The contents of this paper are a condensed version of his thesis. Taylor, his wife Judy, and five children recently moved to the Washington, DC area where he is currently employed with Clark Construction.

Dr. Rollin H. Hotchkiss P.E., D.WRE, F.ASCE, Brigham Young University

Evaluating the Effects of Non-Anonymity on Student Team-Member Evaluations

Introduction

From experience it seems that most students (and people) see the great value of open and honest constructive criticism, but have rarely had opportunities in which they felt safe to provide it. In the procedure which was carried out for this study, students were instructed on how to provide effective feedback, as well as given opportunities to give and receive feedback to their teammates on a group project. This provided them with experiences which increased their self-awareness, as well as taught them how to communicate more effectively. Ideally, this will help them to be more prepared for working on teams in their future careers.

The purpose of this inquiry was to investigate the effects of non-anonymity upon student teammember evaluations, (this is in contrast to the traditional method in which students evaluate one another anonymously). More particularly, this study gauged the effects of non-anonymity upon the self-awareness and defensiveness of those who receive feedback, the willingness of those who provide feedback to be open and honest, as well as its effects upon teams' levels of trust and unity, and levels of performance.

Creating conditions under which people who give feedback will be open and honest, and those who receive feedback will be accepting of it, is essential. This is in order to properly develop the self-awareness of individuals so that they can correctly identify and improve upon their weaknesses. It seems that this is what will lead to an increase in a team's level of trust and unity, and, in turn, its performance.

Note: This paper is a condensed version of the author's Master's Thesis. For a more thorough and comprehensive discussion of the issues surrounding this topic, refer to the original thesis.

Research Questions

- 1. What effect does non-anonymity have upon a recipient's level of self-awareness?
- 2. What effect does non-anonymity have upon the candor and honesty of those who provide feedback?
- 3. What effect does non-anonymity have upon a recipient's level of defensiveness?
- 4. What effect does non-anonymity have upon a team's level of unity and trust?
- 5. What effect does non-anonymity have upon a team's performance?

Literature Review

Self-awareness is the ability which people have to observe and analyze themselves, and is a characteristic which is unique to human beings¹; it is essential to personal growth and development.² Areas that an individual is unaware of personally, yet are known to outside observers, are referred to as *blind spots*¹. It has been well documented that people (on average) hold inflated views of themselves³, and that they tend rate themselves leniently.⁴ Although, one should not automatically assume that this over-optimism results from arrogance; in most instances it is due simply to human limitations and a lack of information.³

The two primary methods for developing self-awareness are self-reflection (internal) and feedback from others (external). Although introspection may provide some self-knowledge, relying too heavily on oneself can be problematic.⁵ Harris articulated this well when he stated:

In order to know oneself, no amount of introspection or self-examination will suffice. You can analyze yourself for weeks, or meditate for months, and you will not get an inch further – any more than you can smell your own breath or laugh when you tickle yourself.⁶

On the other hand, receiving feedback from others is the most effective way to attain self-awareness.^{1,2,3,5} This is because "others" make their observations from a more objective position, and likely have more factual insight.^{3,7}

In terms of self-awareness, it is often unexpected negative feedback that receives the most attention. This is because it is the kind that identifies blind spots. There is an underlying tension that accompanies negative feedback. On one hand, people can gain valuable insights about themselves. On the other hand, exposing blind spots can cause feelings of shame and embarrassment. As a result, people often become defensive and self-justify their weaknesses when they receive negative evaluations.^{1,6,8,9} They may also interpret the feedback as a personal attack, and have a desire to retaliate against those who rated them poorly.¹⁰

With this in mind, the true test of validity for a feedback procedure is *acceptance* by the person being rated. Acceptance is determined by the degree to which the feedback is perceived as being accurate; useful and meaningful; and the extent to which the feedback is used to achieve personal growth. If feedback is interpreted as unfair or biased, it will likely be rejected.^{2,11,12,13} Even if feedback is negative, if a person deems it as relevant and accurate, they are much more likely to internalize it.^{6,14,15}

Due to its sensitive nature, it is also generally quite difficult for people to provide negative feedback to others. As a result, one of the primary biases of the feedback process is leniency. It has been well established that the majority of ratings are positively skewed and that most people receive higher ratings than they actually deserve.^{16,17} What raters may not realize is that despite its uncomfortable nature, open and honest feedback is essential for developing self-awareness and bringing about self-improvement.¹

Supportive communication is the open and honest expression of constructive criticism, which still seeks to maintain goodwill and positive relationships. Due to its honesty, as well as the level of emotional investment required, supportive communication can actually strengthen relationships. Although there may be some initial resistance, people usually end up feeling accepted, valued, and uplifted, even though the information being communicated is negative.⁶

The central focus, and most polarizing issue of this study, is anonymity. Many suggest that in order to achieve truthfulness and candor, and in order to avoid damaging relationships, feedback needs to be given anonymously.^{2,18} On the other hand, one of the drawbacks to anonymity is that there is a lack of accountability. Although anonymous feedback may be less susceptible to leniency bias, the lack of accountability likely diminishes the quality of ratings.¹⁹ It seems that

this is because anonymity does not provoke raters to be as thoughtful and thorough in their responses.

Once again, the whole point of the feedback process is objectivity, and feedback which is not conscientiously fashioned can be just as unrealistic and inaccurate as feedback which is watereddown by leniency. The main benefit to non-anonymity is that it naturally demands justification, which causes the rater to be more thoughtful and deliberate.²⁰ Another major issue with anonymous feedback is that it may send the message, and create an atmosphere where people feel that they cannot trust one another to be honest.²¹ Considerable transparency and trust are required in order for feedback to be given openly.²

Hypothesis

The central hypothesis of this research is that, when students are taught and prepared to properly give and receive constructive criticism, non-anonymous feedback is the most effective. This is based upon the idea that by increasing the level of accountability, the people giving feedback will be more thoughtful and conscientious. This will in turn produce more accurate responses. It also seems that non-anonymity should provide the level of transparency necessary to achieve a high level of interpersonal trust. As a result of higher trust, those receiving the feedback should also perceive the assessments made by others as being more accurate. It seems that if a person views feedback as realistic and unbiased they should be less likely to become defensive, and more likely to internalize it and bring about personal change.

It is important to note that in the feedback procedure which was followed, there were three evaluations performed (as will be discussed hereafter). In many traditional cases students only perform one evaluation at the end of a project. It is supposed that this repetition and follow-up is what should make it possible to overcome the difficulties associated with non-anonymity. It seems that with each successive iteration students should feel more and more comfortable, both giving and receiving feedback.

Subjects

The subjects of this research were students at a large, private, faith-based university in the western United States. They included students enrolled in a sophomore-level leadership class primarily for engineering students (three class sections, 239 students total), as well as seniors enrolled in a civil engineering capstone course (one class section, 52 students). A total of 291 students participated in the feedback procedure, with 239 eligible subjects used in the data set. All of the participants worked on teams for a major term project. The sophomore-level students worked on a non-technical "social change" project, and were for the most part in groups of five. The students in the capstone class worked on an engineering design project, and were generally in teams of four.

Procedure

A feedback procedure was developed for this study which incorporated many of the elements of effective feedback. These elements included: rater training, requiring justification of ratings, time for reflection and goal setting, disclosure, as well as repetition and follow-up.

In preparation for the evaluations students did some reading on supportive communication and received a ten-minute PowerPoint lecture on effective feedback. In order to keep the students blind to the treatment, there was nothing mentioned regarding the pros and cons of anonymity

and non-anonymity. They were also not made aware that other class sections were providing feedback in the opposite manner (i.e. anonymous vs. non-anonymous).

There were three evaluations performed throughout the semester. For each evaluation, the students rated their teammates (as well as themselves) on a scale from 1 to 7 in eight different areas of performance related to team effectiveness. There were also two open-ended questions in which the raters could identify the greatest weakness and strength of the person they were rating, as well as justify and clarify any ratings. The survey was conducted via a web-based survey tool, and the students received a written report of their feedback. The only difference in the format of the reports was whether or not the names of the evaluators were attached to the responses.

Once students received the feedback, they had a chance to review and reflect upon it. They were instructed to then set a personal goal, which they were to share with their teammates at a scheduled team meeting. This same procedure was essentially followed at each of the three time intervals. After the third and final evaluation, the participants also completed a follow-up survey in order to assess the effects which anonymity/non-anonymity had upon their experience.

Experiment Design and Method of Analysis

This study was a random block design, with each class section being randomly assigned to one of the two treatments, anonymous or non-anonymous. There was no treatment effect within class sections because all students within a particular class were subjected to the same conditions.

For the statistical analysis of self-awareness data (i.e., hypothesis 1), hierarchical linear modeling (HLM) was used. HLM was chosen in order to account for variability across the class sections, especially due to the differences in the instructors. This was also necessary due to the capstone class having a completely different curriculum, and because it was composed primarily of seniors. It was also used so that changes in the students over time could be examined (see *Applied Longitudinal Data Analysis* by Singer and Willett).

Self-awareness was measured by a difference score, which represents the algebraic difference between the students' self-rating and the mean of the ratings provided by their teammates. The scores used for analysis were obtained from the "Overall" category of the team-member evaluation. Absolute difference was used because the question of interest was only regarding the ratee's accuracy, and not whether their self-ratings were high or low.

The follow-up survey was designed to provide measures related to the other four primary research questions (i.e., candor and honesty, acceptance, team unity and trust, and team performance), not directly related to self-awareness. Within the survey there were multiple questions which were intended to measure each of these constructs, and Cronbach's alpha was used to determine their correlation within each construct. T-tests, cross-tabulation, and odds ratios were then used to compare the differences in the responses between the two groups.

Limitations

For the analysis of self-other agreement, one of the limitations to the data is that not all students completed the evaluations. As a result, in some instances the mean other-score used in the analysis was based upon only 1 or 2 raters.

For the follow-up survey, for each construct (i.e., trust, performance, etc.) a mean score was calculated for each team. As a result, although the sample size for the individual students was quite adequate, the sample size for the number of teams was not very large ($N_{Anon} = 33$, $N_{Non} = 32$). In addition, team performance was measured from student responses on this survey. It would have been more preferable to use another variable, such as grades or some other standardized measure, to evaluate performance more explicitly; however, due to the variability in grading across sections, as well as due to the fact that the capstone students completed an entirely different project, this was not deemed a viable option.

Results

In this section, only a limited number of results will be discussed. It is suggested that the reader refer to the original thesis for a more comprehensive discussion of the results.

Hypothesis 1: Self-other Agreement

The feedback procedure which students followed did increase their level of self-other agreement (p=0.000), although there was no significant difference between the anonymous and nonanonymous groups. However, a plot of the actual results provides some interesting insights. Figure shows the absolute difference scores at each evaluation for all participants.





In this figure, the downward trend in both groups illustrates the reduction in the difference between self- and other-ratings, and thus an increase in self-other agreement. It seems that both groups begin and end at approximately the same level of self-other agreement, but for the nonanonymous students it looks as if there is a delayed reaction followed by a period of more rapid progress.

Hypothesis 2: Leniency and Honesty

The analysis showed that over the course of the feedback procedure (i.e. longitudinally) there was no significant difference between the treatment groups (p-value 0.467); however, a t-test for difference of means was performed for each time interval. The actual scores provided at each interval are plotted in Figure 2, with the statistical analysis shown in 1.



Figure 2: Plot of Other-scores at Evaluations 1, 2, and 3

Evaluation #	Mean _{Anon}	Mean _{Non}	Difference	p-value
1	6.47	6.55	0.08	0.105
2	6.60	6.50	-0.10	0.047
3	6.75	6.72	-0.03	0.330

Table 1: Statistical Analysis of Other-scores at Evaluations 1, 2, and 3

At time-1, the non-anonymous scores were higher than the anonymous, and the p-value of 0.105 provides suggestive but inconclusive evidence that the non-anonymous ratings were more lenient at the beginning of the procedure. By looking at the beginning and end time intervals, the respective p-values of 0.105 and 0.330 seem to confirm the hypothesis that non-anonymous other-scores start out higher and then by time-3 there is no difference; however, what is most interesting is that time-2 produced the most statistically significant difference, and that the non-anonymous scores were actually lower than the anonymous scores. This raises a question as to why the most significant difference came in the middle of the procedure. Furthermore, was this drop due to a decrease in performance, or did something cause raters to grade more harshly? Upon further reflection, it seems that that the drop resulted not from decreased performance, but rather from non-anonymous raters becoming more comfortable providing more honest and accurate feedback (due to the fact that they are more susceptible to leniency).

In addition to the ratings themselves, the follow-up survey was also used to measure honesty and candor. The questions were designed as a gauge of the students' perception of the overall process, rather than at particular points of time. Table 2 displays these results.

Survey Question	Mean _{Anon}	Mean _{Non}	p-value
How honest and candid students were students when providing positive feedback.	4.68	4.54	NS
How much more open they would have been providing positive feedback if they were in the other treatment group.	3.20	3.28	NS
How honest and candid students were students when providing negative feedback.	4.10	3.67	< 0.004
How much more open they would have been providing negative feedback if they were in the opposite treatment group.	2.54	3.55	<0.001

Table 2: The Honesty and Candor of Feedback

When it came to providing positive feedback, the treatment had no effect upon honesty. However, for negative feedback the p-values <0.004 and <0.001 provide very convincing evidence that non-anonymity has a negative effect on students' ability and/or willingness to be honest. Once again, these conclusions are regarding students' overall perceptions, and not in regards to one type of feedback or one specific point in time.

Hypothesis 3: Defensiveness and Acceptance

The results of t-tests for the analysis of the thoughtfulness of the rater, as well the level of acceptance by the ratee, are shown in Table .

Survey Question	Mean _{Anon}	Mean _{Non}	p-value
How much time and thought they invested when providing feedback.	3.80	3.77	NS
Whether the treatment made them more or less defensive towards negative feedback.	2.83	3.11	< 0.03
How much they valued the feedback they received.	4.17	4.06	NS
The extent to which they regarded the feedback as sincere and thoughtful.	3.87	4.01	NS

Table 3: Thoughtfulness and Acceptance of Feedback

According to this analysis, there was no treatment effect regarding how much time and thought raters invested in providing feedback, how much people valued the feedback they received, as well as the extent to which they regarded it as being thoughtful and sincere. However, there was an observable difference regarding defensiveness. The p-value of <0.03 provides moderate evidence that, contrary to the hypothesis, those who receive feedback non-anonymously feel more defensive than those who receive it anonymously.

Hypotheses 4 and 5: Team Cohesiveness and Performance

Just as with defensiveness and acceptance, team cohesiveness and performance were gauged by the students' perceptions (as discussed previously). The analyses showed that the treatment had no effect upon a team's cohesiveness, as well as its performance. As such, there will be no discussion provided herein.

Hypothesis 6: Future Choice

One of the final questions in the survey asked the students which method of feedback, anonymous or non-anonymous, they would choose in the future. The two-way table for the chisquare analysis is provided in Table below.

		Type of group they were in		
		Anon.	Non-Anon.	Total
Type of group they would choose in the future	Anon.	104 (75.4%)	44 (41.5%)	148 (60.7%)
	Non-Anon.	34 (24.6%)	62 (58.5%)	96 (39.3%)

Table 4: Two-way Table for Treatment Group vs. Future Choice

This two-way table is statistically significant, with a chi-square p-value <0.001, indicating conclusively that there is a relationship between the group that the students were in and the method they would choose in the future. Of those in the anonymous group, 75% would choose to be in an anonymous (same) group again in the future, and 25% would choose to be in a non-anonymous (different) group. For non-anonymous, 42% would choose to switch to anonymous (different), and 58% would stay non-anonymous (same). Therefore, within each group, a majority of participants would prefer to stay with the same method, although those in the anonymous group show a stronger preference.

In addition, the odds of anonymous participants choosing non-anonymous in the future are 0.32 (34/104 = 0.32) and for the non-anonymous the odds are 1.41 (62/44=1.41). The odds ratio of 0.23 (0.32/1.41=0.23) indicates that those in the anonymous group are 77% less likely than those in the non-anonymous group to choose to provide feedback non-anonymously in the future. In other words, the non-anonymous group is 4.41 times more likely to choose non-anonymous feedback in the future than is the anonymous group. In contrast, the anonymous group is 4.31 times more likely than the non-anonymous group to choose anonymous feedback in the future ((104/34)/(44/62) = 4.32).

Summary

Table 5 provides a summary of the results.

Table 5: Summary of Results

Hypothesis 1:	Non-anonymous feedback will result in a greater increase in self-other agreement than the anonymous students.
Conclusions:	The procedure which students followed increased their level of self-other agreement (p=0.000), although there was no significant difference between the anonymous and non-anonymous groups. There was also no significant difference in the change in self- and other-scores.
Hypothesis 2:	In the first evaluation, other-scores will be higher for the non-anonymous students, but there will be no difference at evaluation #3. There will also be no difference, as perceived by the students, in the level of honesty and candor between the groups.
Conclusions:	At time-1 the non-anonymous other-scores were somewhat more lenient than the anonymous other-scores (p=0.105), but by time-3 there was no significant difference (p=0.330). At time 2 there was a significant difference (p=0.047) between the groups, and the non-anonymous scores were lower. Students who provided feedback non-anonymously were less honest and candid (p<0.004), and they perceived the treatment effect to have had a greater influence upon their ability/willingness to be honest than the anonymous students (p<0.001).
Hypothesis 3:	Non-anonymity will produce more thoughtful, sincere, and accurate feedback, and those receiving it will be less defensive.
Conclusions:	There was no treatment effect regarding how much time and thought that raters put into feedback, how much people receiving feedback valued it, and also the extent to which they regarded it as being thoughtful and sincere; however, those who receive feedback non-anonymously feel more defensive $(p<0.03)$.
Hypothesis 4:	The non-anonymous students will be more comfortable expressing opinions, as well as dealing with conflicts. These teams will also experience higher levels of cohesiveness.
Hypothesis 5:	The performance of non-anonymous teams will be rated higher than that of the anonymous teams.
Conclusions:	There was no treatment effect upon team cohesiveness and performance.
Hypothesis 6:	Students in the non-anonymous group will be just as likely to prefer non- anonymous feedback in the future as those in the anonymous group will be to prefer anonymous in the future.
Conclusions:	There is a relationship between the group that the students were in and the method which they would choose in the future. Non-anonymous are 4.41 times more likely than anonymous to choose non-anonymous in future. Anonymous are 4.31 times more likely than non-anonymous to choose anonymous in the future. Thus the odds are roughly the same.

Additional Conclusions from Free Responses

Taking an overall look at the analyses, there were only a few significant differences between the two groups. Even then, some of these differences were in support of the hypothesis, while others were not. In addition, the comments from the follow-up survey were generally balanced as to the pros and cons of anonymity versus non-anonymity, supporting the notion that anonymity and non-anonymity were not prime determinants of the effectiveness of the feedback procedure. Many of the other responses to the follow-up survey provided additional and valuable insights into why this was so.

First, the most common response which students provided was that the evaluations were not adequately spaced. (This process took place over about a three-month period, and feedback was received only 3-4 weeks apart). Although students saw the value of providing feedback multiple times, many felt that this was too short of a period for team members to bring about any observable changes in behavior. This seems to suggest that if a treatment effect does in fact exist, it may take long periods of time for these differences to be realized. Secondly, it seems that with small groups it is difficult to maintain anonymity. Many of the anonymous participants stated that they were able to figure out who it was that provided them with feedback, essentially removing the treatment.

Recommendations

Upon considering the material covered in the literature review, as well as insights which were gained from the follow-up survey, the following recommendations for improvement to the procedure are being made, (keeping in mind that some of the procedures followed were necessary in order to obtain the data for analysis).

One of the benefits to an early first evaluation is that it gives students more "practice", which makes them better prepared to provide higher quality feedback on subsequent evaluations; however, some students suggested that the first evaluation was less effective because they felt it was too early in the project to be able to provide accurate and useful results. One of the key components of a successful feedback procedure is maintaining the credibility of the process. If participants view the procedure with skepticism and contempt, then the process is essentially rendered useless.

In order to address this issue, it is recommended that, rather than performing a team member evaluation at time-1, students should perform a 360-feedback survey. In a 360-analysis, students rate themselves on different leadership and interpersonal competencies, and their other-ratings come from stakeholders in other areas of their lives (not their teammates). The team would still hold a meeting early in the project in which the team members openly declare their goals for the project to the team; however, the goals would be based upon the results of the 360-feedback rather than feedback from their teammates. The current procedure would then resume for evaluations 2 and 3. This modified procedure would still accomplish the purposes of increasing self-awareness, and establishing patterns of communication and trust, but without raters having to perform an extra evaluation. In addition to avoiding skepticism, avoiding rater burnout is critical to maintaining the credibility of the procedure.

Another way to reduce the burden on raters is to eliminate the self-ratings at times 2 and 3. The 360-feedback at evaluation-1 would already include a formal self-evaluation and would achieve

the purposes of heightening self-awareness. Even though students would not formally rate themselves for evaluations 2 and 3, when people receive ratings and feedback from others they instinctively "rate" themselves internally; it's a part of human nature. It seems that with subsequent evaluations, formal self-ratings become redundant and are viewed simply as a task to be performed.

It would be beneficial to have students administer their own feedback, although they would all still be required to follow a uniform procedure. There are multiple reasons why this is beneficial. First, it ensures confidentiality. If a student thinks that those administering the feedback (i.e., TA or professor) are going to view the results, it may skew their responses. Self-administration is also beneficial because it creates a greater sense of ownership for the feedback; that is, it seems that because students would have more responsibility and control, and would thus take the process more seriously. Third, self-administration simplifies the process and reduces the potential for error. Each individual "administrator" would only be responsible for managing 3 or 4 evaluations, whereas when one person administers the feedback for all participants he or she is potentially responsible for hundreds of responses.

For the open-ended feedback, the two questions regarding "strengths" and "weaknesses" should be combined into one question in which students are free to address either one or the other, or both. With the way that these questions were structured, some students felt obligated to find and report weaknesses in their teammate, even if they felt that none existed, thus stifling openness and accuracy. In addition, at the recommendation of one of the students in the follow-up survey, rather than "weaknesses", a term such as "areas of most needed improvement" should be used. This will likely make raters more willing to address areas of concerns, as they will not feel as punitive in providing negative feedback.

One final suggestion is to ensure that the rating scale is simple. One student commented, "The numbering system was strange. What does a 6.8 out of 7 mean? Well it probably means the same thing as a 4 in a 1 to 5 scale. ... Don't make it so complicated." In this procedure the students were rated on a scale of 1 to 7, although the scale increased at increments of 0.1, essentially creating a 70 point scale. A full integer or half-point scale would be more appropriate.

Conclusions

Despite any shortcomings in the process, the overall impression from the follow-up survey was that this feedback procedure was effective at achieving its primary objective which, as stated in the introduction, was to increase self-awareness in students and "provide them with experiences which will teach them how to communicate more effectively and prepare them for working on teams in their future careers." The following is a sample of the comments which were received from students which validate this claim.

This was a great experience. Personally, I want to use this idea when I am a head of a group ... It really helped our team head for success.

I thought that providing feedback was helpful in allowing my group to work better together. We were able to understand what everyone was doing and wanting and it was helpful to know where I could improve based on the thoughts of my group members. It was hard for me to give non-anonymous feedback. But it caused me to be more thoughtful and sincere about critiquing my teammates. I think that we need to get used to giving non-anonymous feedback, because that's how it is in most real-life situations. Setting specific goals based on feedback and sharing them with the team was particularly helpful to my progress.

I felt that as I completed the final feedback for my group I had finally developed the ability to be open, honest, and candid with the team members. ... Overall I am very grateful for the opportunity I had to participate in the feedback three times. It was great! Thanks for the experience. It really helped me come to trust my teammates and their opinions.

I think that this is an effective way of improving team cohesiveness and ability. I really enjoyed seeing the feedback, and understanding what my team wanted or was requiring of me. I also think that this was even more effective because I was team leader.

It was good. We all learned vital lessons from each other. We dealt with conflict well, and [I'm] excited to talk to them in the halls later.

From a statistical perspective, there was no strong evidence for or against non-anonymity, and thus it appears that there was no major treatment effect. The primary conclusion from this research is that it appears that, under the conditions which were present, it may be difficult to obtain statistically significant differences between anonymous and non-anonymous feedback. It seems that this is due to two primary factors. First, if non-anonymous feedback does indeed produce positive outcomes, it may take long periods of time for these differences to be noticed. Secondly, when teams are small (i.e., only 3-5 members) it is difficult to maintain anonymity. If anonymity is absolutely essential, perhaps this is only possible in larger groups.

Perhaps the most delicate issue with peer feedback is establishing an appropriate level of accountability under which people will give impartial, thoughtful, and candid feedback, and those who receive it are accepting of it and implement it in their lives.

One proposed solution to this problem is to utilize either anonymous or non-anonymous feedback at different stages of the process. In the early stages, confidentiality and anonymity may be essential for instilling confidence in the feedback process; however, once people have become accustomed to the procedure, a feedback system which includes the open discussion of ratings could be adopted.²²

Alternatively, the preferred procedure is for students working in small teams to provide nonanonymous feedback throughout the entire process. This is because, as just noted, even if anonymity is a required condition, it is difficult to maintain. It seems that pretending that anonymity exists, when in fact it does not, may actually hinder transparency and trust. Furthermore, even if students are unsure of whom it is that provided them with feedback, there still exists the natural tendency to make assumptions about who it was, which can lead to false accusations (although they may never be spoken out loud), and increase tension and mistrust within a team. Finally, it seems that giving feedback non-anonymously will more effectively prepare students for working on teams in their careers, as this is more reflective of the way that feedback is provided in the workplace. In practice this second proposal may result in some initial resistance from participants; however, requiring non-anonymity from the beginning will accelerate their transition to providing open and honest feedback. Either way, the highest aim of the feedback process should be the cultivation of an atmosphere of transparency and trust, in which both the giver and receiver are fully accountable, and opinions and feelings are shared openly and freely. When this occurs, formal feedback and questionnaires eventually disappear and are replaced with ongoing feedback on an informal, day-to-day basis.^{2,23}

REFERENCES

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3208397/pdf/fpsyg-02-00312.pdf>.

- ⁷ Fleenor, J. W., Smither, J. W., Atwater, L. E., Braddy, P. W., and Sturm, R. E. (2010). "Self-other rating agreement in leadership: A review." *The Leadership Quarterly*, 21, 1005–1034.
- ⁸ Delmhorst, F. P. (2006). "Self-awareness and leadership development: Examining the impact of self-subordinate agreement on perceptions of leadership." Ph.D. dissertation, Columbia University, New York, NY.
- ⁹ Baumeister, R. F., and Cairns, K. J. (1992). "Repression and self-presentation: When audiences interfere with self-deceptive strategies." *Journal of Personality and Social Psychology*, 62(5), 851-862.
- ¹⁰ Taggar, S. and Brown, T. C. (2006). "Interpersonal affect and peer rating bias in teams." *Small Group Research*, 37(1), 86-111.
- ¹¹ Ilgen, D. R., Fisher, C. D., and Taylor. M. S. (1979). "Consequences of individual feedback on behavior in organizations." *Journal of Applied Psychology*, 64, 349-371.
- ¹² Bracken, D. W., and Timmreck, C. W. (2001). "Success and sustainability." *The handbook of multisource feedback: The comprehensive resource for designing and implementing MSF processes*, D. W. Bracken, C. W. Timmreck, A. H. Church, eds., Jossey-Bass, San Francisco, CA, 478-494.
- ¹³ Murphy, K. R., Cleveland, J. N., and Mohler, C. J. (2001). "Reliability, validity, and meaningfulness of multisource ratings." *The handbook of multisource feedback: The comprehensive resource for designing and implementing MSF processes*, D. W. Bracken, C. W. Timmreck, A. H. Church, eds., Jossey-Bass, San Francisco, CA, 130-148.
- ¹⁴ Bracken, D. W., and Rose, D. S. (2011). "When does 360-degree feedback create behavior change? And how would we know it when it does?" *Journal of Business Psychology*, 26, 183-192.
- ¹⁵ Peterson, D. B., Hicks, M. D., and Stoner, J. D. (2001). "Tools and resources for helping people move forward following multisource feedback." *The handbook of multisource feedback: The comprehensive resource for*

¹ Luft, L. (1969). *Of Human Interaction*, National Press, Palo Alto, CA.

² Tornow, W. W., and Paradise, C. (2001). "Linking multisource feedback content with organizational needs." *The handbook of multisource feedback: The comprehensive resource for designing and implementing MSF processes*, D. W. Bracken, C. W. Timmreck, A. H. Church, eds., Jossey-Bass, San Francisco, CA, 48-62.

³ Dunning, D. (2005). *self-insight: roadblocks and detours on the path to knowing thyself*, Psychology Press, New York, NY.

⁴ Atwater, L. and Yammarino, F. (1992). "Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions?" *Personnel Psychology*, 45.

 ⁵ Bollich, K. L., Johannet, P. M., and Vazire, S. (2011). "In search of our true selves: Feedback as a path to self-knowledge." *Frontiers in Psychology*, 2, 312,

⁶ Whetten, D. A., and Cameron, K. S. (2007). *Developing Management Skills*, 7th Ed., Pearson Education, Upper Saddle River, NJ.

designing and implementing MSF processes, D. W. Bracken, C. W. Timmreck, A. H. Church, eds., Jossey-Bass, San Francisco, CA, 221-238.

- ¹⁶ Murphy, K. R., and Cleveland, J. N. (1995). Understanding performance appraisal: Social, organizational, and goal-based perspectives, Sage, Thousand Oaks, CA.
 ¹⁷ Demonstrational and the second second
- ¹⁷ Bernardin, H. J., Cooke, D. K., and Villanova, P. (2000). "Conscientiousness and agreeableness as predictors of rating leniency." *Journal of Applied Psychology*, 85(2), 232-234.
- ¹⁸ Antonioni, D. (1994). "The effect of feedback accountability on upward appraisal ratings." *Personnel Psychology*, 47(2).
- ¹⁹ Cleveland, J. N., and Murphy, K. R. (1992). "Analyzing performance appraisal as goal-directed behavior." *Research in Personnel and Human Resources Management* (Vol. 10), G. Ferris and K. Rowland, eds., JAI Press, Greenwich, CT.
- ²⁰ Ford, J. K., and Weldon, E. (1981). "Forewarning and accountability: Effects on memory-based interpersonal judgments." *Personality and Social Psychology Bulletin*, 7, 264-268.
 ²¹ Mitchell, T. W. and Williem and Social Psychology Bulletin, 7, 264-268.
- ²¹ Mitchell, T. W., and Klimoski, R. (1984). "Accountability bias in performance appraisal." *Working Paper Series in Industrial/Organizational Psychology*, Ohio State University, Columbus, OH.
- ²² Peiperl, M. A. (1999). "Conditions for the success of peer evaluation." *International Journal of Human Resource Management*, 10, 429-458.
- ²³ Antonioni, D. (1996). "Designing an effective 360-degree appraisal feedback process." Organizational Dynamics, 25(2), 24-38.