

3-23-2015

How do musicians evaluate their musical performances? The impact of positive and negative information from normative, ipsative, and expectation standards

Ellen-ge Denton
CUNY College of Staten Island

William F. Chaplin
St. John's University

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: https://academicworks.cuny.edu/si_pubs

 Part of the [Psychology Commons](#)

Recommended Citation

Denton, Ellen-ge and Chaplin, William F., "How do musicians evaluate their musical performances? The impact of positive and negative information from normative, ipsative, and expectation standards" (2015). *CUNY Academic Works*.
https://academicworks.cuny.edu/si_pubs/99

This Article is brought to you for free and open access by the College of Staten Island at CUNY Academic Works. It has been accepted for inclusion in Publications and Research by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.

How do musicians evaluate their musical performances? The impact of positive and negative information from normative, ipsative, and expectation standards

Psychology of Music

1–14

© The Author(s) 2015

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0305735614568883

pom.sagepub.com**Ellen-ge Denton¹ and William F. Chaplin²****Abstract**

The purpose of the research reported in this article was to test two hypotheses about how musicians evaluate their musical performances. The first hypothesis was that musicians' self-evaluations would be more influenced by their expectations and their past performances than by comparisons to the performances of other musicians. The second hypothesis was that musicians would exhibit an 'adaptive evaluational style' by showing more sensitivity to positive feedback than to negative feedback. We used the Experimental Evaluational Styles Questionnaire (Goolsby & Chaplin, 1988) in a sample of 78 music performance students (43 men and 35 women) to test these hypotheses, and both were supported. These results represent one of the first examples where the dominant theory of evaluation in psychology, Festinger's (1954) social comparison theory, did not have the greatest influence on people's performance evaluations. However, we did find individual differences in the influence of the different evaluative standards. Understanding the causes and consequences of these individual differences should be a fruitful target for future research.

Keywords*competition, evaluation, musical performance, social comparison*¹CUNY – Staten Island, USA²St. John's University, USA**Corresponding author:**

Ellen-ge Denton, Department of Psychology, City University of New York College of Staten Island, Room 4S-229, 2800 Victory Boulevard, Staten Island, NY 10314, USA.

Email: ellenge.denton@csi.cuny.edu

Competitions are for horses, not artists.

Béla Bartók (1881–1945)

The dominant theory in psychology of how people evaluate their experiences, accomplishments and performances is Festinger's (1954) Social Comparison Theory. According to this theory, people use the normative standard of comparing to others to evaluate performances for which no objective, absolute standard exists. Although there is a vast literature supporting the power of social comparison, far less research has evaluated social comparison processes in relation to other standards that are less competitive and more self-focused, such as the ipsative standard of improvement or change over time, and the expectation standard based on goals and ideals. In this research we specifically selected a domain, artistic musical performance, where performers generally deny comparison to others in evaluating their work, making musical performance a likely domain for ipsative and expectation standards to dominate. We also considered how musicians respond to positive (downward) and negative (upward) evaluations of their musical performances. In this research, our focus is on performing musicians and we restrict the term 'musician' to this group.

Standards of comparative evaluation

Social comparison theory (Festinger, 1954) has dominated research on self-evaluation since it was proposed. Although the theory has evolved over the years, the basic tenet is that people evaluate their performances, experiences, and behavior by comparing them to those of others. There is little doubt that people do engage in social comparisons, and external standards have an impact on how people evaluate themselves (e.g., Suls & Wheeler, 2000; Suls & Wills, 1991; Taylor, Wayment, & Carrillo, 1996). Indeed, social comparison has so dominated the literature on self-evaluation that it is automatically invoked as the basis for interpreting scores on clinical, ability, and personality measures through the use of 'norms.'

However, although the normative standard of comparison to others has dominated the self-evaluation literature, other standards are recognized and have been used. One such standard was described by Albert (1977) as 'temporal comparison theory' as a direct counterpoint to social comparison theory. In temporal comparisons, the focus of evaluation is on the self rather than others with an emphasis on how one's performances and behaviors have changed over time. In the language of modern psychometrics, such a standard is called 'ipsative.' Using an ipsative standard, if one's performance is improving one is doing well, and it matters not how that performance compares to the performance of another person.

A third broad type of evaluative standard is based on comparisons to ideals, goals, or our expectations. Such standards are central to Higgins' (1987) self-discrepancy theory, and in contrast to normative or ipsative standards, such ideal or expectation standards do not require an external reference either to other people or to one's past. Instead such standards can be imaginary as in Markus and Nurius's (1986) 'possible selves.' Lamiell (1981) discussed the relevance of such idealized standards, which he dubbed 'idiotic,' for personality assessment.

One other important aspect of social comparison theory is the direction of the comparison. Conventionally, the direction of comparison is labeled either upward or downward. In the context of performance evaluations, downward comparisons are generally positive evaluations ('performed better than . . .') whereas upward comparisons are generally negative ('performed worse than . . .'). Therefore, an individual may assess his or her relative standing to be 'higher' or 'better than' normative, ipsative, and idiotic standards (downward comparison) or an individual may assess his or her relative standing to be the opposite, 'worse than . . .' on comparison

standards. Obviously, the direction of comparison has different effects on evaluation, such that downward comparisons lead, in this case, to positive feelings, and upward comparisons lead to negative feelings (Aspinwall & Taylor, 1993). In addition, upward and downward comparisons can serve different self-regulation goals, such that downward comparisons may enhance self-esteem and upward comparison may serve as motivations to improve (Langer, Delizonna, & Pirson, 2010).

Disentangling the effects of the different standards on evaluations

As established in the previous section, when considering the evaluation of a musical performance, it is possible to base the evaluation on comparisons to others' performances (normative), comparisons to one's own past performances (ipsative), and/or an idealized performance (expectation). A major challenge in isolating the effects of each standard on a person's evaluation is that these three standards tend to be highly correlated (Chaplin & Buckner, 1988). Thus, disentangling their separate influences cannot be done naturalistically. Goolsby and Chaplin (1988) developed a method for experimentally separating the effects of information based on each of the three standards on individual's evaluations. The method is called the Experimental Evaluational Styles Questionnaire (EESQ), because it represents, in questionnaire form, a fully crossed $3 \times 3 \times 3$ repeated measures experimental design in which each of the three standards (normative, ipsative, expectation) has three levels that represent downward (positive), upward (negative), and neutral (absent) comparative information about a particular event or performance. The specific version of the EESQ for evaluating a musical performance (EESQ-M) is described in the Methods section below.

The logic of this method is that it contrasts the degree of negative feelings about a performance when a particular standard is upward ('worse than') with the degree of how positive a person feels when a particular standard is downward ('better than'). The difference between the average upward rating and the average downward rating of a standard is in the context of all possible combinations of the other two standards, thus experimentally isolating the standard of interest. This assessment can be done by testing the main effects of the normative, ipsative, and expectation evaluative standards in a $3 \times 3 \times 3$ repeated measures analysis of variance. One can also compute the average rating of the 9 items that contain the upward version of a standard and subtract that average from the 9 items that contain the downward version of that standard to obtain an individual difference measure of that standard's influence. The larger the obtained difference between upward and downward directions of the standard, the more influential the standard, while controlling for all other standards. For example, a difference of zero would imply no influence of that standard because the individual responds the same regardless of whether the standard is upward or downward. In contrast, a difference of 6 (when using a 1-to-7 scale) would imply maximum influence, as the person feels terrible in response to an upward comparison and ecstatic in response to a downward comparison.

The dependent variable in this design is a simple affect rating of 'How would you feel (when receiving this information)?' For example, one item might state 'You performed better than other people, but worse than you have in the past and worse than you expected to perform . . . How do you feel?' Ratings range from 1 (extremely bad) to 7 (extremely good). We chose affect as our dependent variable because there is considerable evidence to suggest that peoples' reports of how they would feel about an event in the context of evaluative information can be made quickly, confidently, and accurately (Zajonc, 1980). On the other hand, people are generally very poor at saying what standard they use or at explaining why particular information makes them feel good or bad (Nisbett & Wilson, 1977). Thus, we did not ask the participants to directly rate how much a standard influences them as they likely would not know. Instead, we gave them evaluative information based on different standards and asked them how they would feel,

which people can do accurately, even when the situation is hypothetical (McArthur, 1972). Based on the differences in the affect ratings, we infer the impact of a particular standard.

Musicians' evaluation of their musical performance

As the quotation that begins this article implies, Bartók does not believe that musicians are primarily influenced by the competitive normative standard of social comparison. Indeed, Bartók suggests that musicians are not at all competitive or normative. These views are also expressed in the modern music literature. For example, Thompson & Williamon (2003) describe musicians as developing an 'internal marking scheme' that is non-specific to the musical piece. These internal schemes may be difficult to express verbally and do not appear to relate to other musicians' performances (Coimbra, Davidson, & Kokotsaki, 2001). In terms of standards of evaluations, these internal grades seem more ipsative or expectation-based. Indeed, Coimbra, Davidson, and Kokotsaki (2001) specifically suggest that musicians work toward an 'ideal state,' which they describe as processing musical work in 'perfect synergy that would result in one energized whole.' McCormick and McPherson (2003) found that musicians sought to 'fully master the requirements for examination,' which seems to also be an expectation-based standard, and Ritchie and Williamon (2011) suggested that conservatoire musicians are known to strive to 'reach professional standards' after entering at an accomplished level.

Overall, authors who have discussed musicians' self-evaluations of their performances generally refer to standards that are idealized or goal-based, or perhaps based on an internal or ipsative standard. But consistent with Bartók's quotation, none of these authors include a social comparison standard in their discussion. Indeed, Priest et al. (2006) specifically suggested that normative comparisons would undermine the creative process and create 'ego- rather than task-involved' musicians. However, the previous literature on musicians' self-evaluations is clearly limited and does not involve any empirical comparisons of the different standards. Also, in contrast to the claims of musicians, the music industry itself is highly competitive, involving auditions, talent searches, music charts, and selling albums. In high school and college bands and orchestras, individual players are comparatively ranked as 'first chair,' 'second chair,' and so on, and, within an academic setting, music performance is a special case of more general academic performance. In the one study that used direct empirical comparisons between normative, ipsative, and expectation standards in the self-evaluation of academic performance (Goolsby & Chaplin, 1988), it was the normative standard that had the strongest impact on student's evaluative ratings.

Finally, as our earlier review suggested people are often very poor at knowing what standards are actually influencing their evaluations. We suspect that people may deny using a normative standard simply because appearing to be competitive in an artistic domain is not socially acceptable. Thus, the present study was designed to provide a direct comparison among the evaluative standards.

Our research design will also allow us to assess the extent to which musicians are more sensitive to downward (positive) or upward (negative) information. Given the competitiveness of the music industry, we suspect that musicians will experience many failures and disappointments before they experience continued success. Of course, music is not unique in this respect as people in many fields and careers have to endure failures on the way to success. Nonetheless, it is the ability to exhibit resilience in the face of negative feedback that is usually considered the key to success (Knudson, 2013). Thus, we suspect that musicians will generally tend to be less sensitive to negative (upward comparison feedback) and will be more influenced by positive (downward comparison) feedback.

The present study

In the present study, we sought to empirically compare the impact that different evaluative standards have on musicians' evaluations of their musical performances. To do this, we used the EESQ adapted for evaluating a musical performance, to disentangle the influence of normative, ipsative, and expectation standards, as well as compare the impact of positive (downward) and negative (upward) information on musicians' performances. We conducted our study on a sample of music students who were enrolled in music performance classes so that our sample would be comparable to the one studied by Goolsby and Chaplin (1988) on general academic performance evaluations. We tested two hypotheses:

- 1) In contrast to general academic performance evaluations, music students will be least influenced by normative comparisons. Based on the literature, we also expect that expectation standards will be the most influential.
- 2) Musicians will respond differently to upward and downward evaluative comparisons, with downward comparisons generally leading to feeling good and upward comparisons generally leading to feeling bad. However, we expect there will be an asymmetry in these responses with musicians showing less sensitivity to negative as compared to positive evaluations.

Method

Participants

Participants were 82 students recruited from the University of Alabama Music Performance Program. However, four students responded with the same rating on all 27 of the items and are not included in these analyses. The remaining 78 students (43 (55%) men and 35 (45%) women) had a mean age of 20.87 ($SD = 4.81$) years. The participants reported 12.63 years of musical experience (Mean; $SD = 6.28$), a mean average of 4.39 ($SD = 1.81$) days per week designated for musical practice with each practice session lasting, on average 2.01 hours (Mean; $SD = 2.13$).

Measures

Experimental Evaluation Styles Questionnaire-Music version (EESQ-M). The primary measure for this study is the EESQ-M, which is a variation of the EESQ developed by Goolsby and Chaplin (1988). This questionnaire consists of 27 items that represent a completely balanced $3 \times 3 \times 3$ experimental design, hence the word 'experimental' in the title. The three factors represent the normative, ipsative, and expectation evaluative standards and the three levels within each standard represent positive (downward comparison), negative (upward comparison) and neutral (absent) information. Thus each item contains either upward (in this case unfavorable), downward (in this case favorable), or neutral information from each standard.

The general instructions and rating scale along with the items on the EESQ-M are shown in the Appendix. As an overview, participants were told that they would be evaluating a musical performance based on information from three different comparative standards. They were told that they would be making 27 different evaluations on a rating scale that concerned their general feelings about the performance given the evaluative information. The scale itself ranges from feeling very bad (1) to feeling very good (7).

The completely balanced design of the EESQ-M allows us to disentangle several aspects of musician's evaluations of their performances. First, the average of the 27 items indicates a person's general evaluation of giving a musical performance across all combinations of evaluative information. The 7-item downward scale indicates how musicians react to only positive information about their performance whereas the 7-item upward scale indicates the reaction to only negative information. The 12-item 'Mixed' scale indicates the reaction when the evaluative information is a mixture of positive and negative.

In addition to these four evaluative scales, it is also possible to derive three Influence scales that indicate the relative impact that normative, ipsative, and expectation scales have on a person's evaluations. These scales are derived by calculating the mean of the nine downward (positive) items for a standard (e.g., normative) and the mean for the nine upward (negative) items for a standard (e.g., normative). The difference (normative downward-normative upward) reflects the influence of normative information, experimentally controlling for the other standard impacts, on a person's evaluations. For example, a score of zero means that upward and downward normative information led to the same average evaluation and hence the normative information had no influence on the person. In contrast, a score of 6 (the maximum possible) would mean all the downward normative items elicited a response of 7 whereas all the upward items elicited a response of 1, *controlling for the other standards*, suggesting maximum normative influence. It is possible for a person to receive a negative score on these scales if a standard has no influence and one of the other standards dominates it. Conventionally, such rare, but possible negative, scores are set to zero.

Personal data. The students provided information on their age and sex as well as information about their musical experience (in years), frequency of practice in days/week, and length of practice in hours/session.

Procedure

Participants were recruited from music performance classes that included both voice and instrument. Such classes are typically available only to music majors. The participants were informed that the study concerned how musicians evaluate their own musical performance, that participation was voluntary and was not part of their class and would have no impact on their grade, and that they would receive no compensation for participating. Generally all students who were in attendance that day agreed to participate, signed the informed consent form, and completed the two questionnaires at that time.

Results

Figure 1 summarizes the mean responses to the 27 items (conditions) from the $3 \times 3 \times 3$ design. As shown in Figure 1 the seven items that contain upward (negative, 'worse than') comparison information elicited the most negative responses whereas the downward (positive, 'better than') comparison information elicited the most favorable responses. The 12 items that contain mixed evaluative information are in the middle. These results might be viewed as a manipulation check on the EESQ-M as the ordering of the items suggests the participants read and understood the questions. Interestingly, the completely neutral item (nnn) elicited a mean response that is well above the neutral point ('4') on the response scale, suggesting that musicians generally feel positive about giving musical performances.

Table 1 summarizes the results of the $3 \times 3 \times 3$ repeated measures MANOVA applied to this design. As shown in Table 1, all the effects from this design are statistically significant, indicating

	Item	Mean
Upward Evaluative Information	uuu	1.93
	uun	2.40
	nuu	2.45
	unu	2.45
	nnu	2.46
	nun	2.72
	unn	2.90
Mixed Evaluative Information	duu	3.34
	udu	3.37
	uud	3.57
	dnu	3.80
	ndu	3.84
	dun	3.99
	udn	4.24
	nud	4.24
	und	4.35
ddu	4.68	
dud	4.70	
udd	4.93	
Neutral item	nnn	5.48
Downward Evaluative Information	dnn	5.35
	nnd	5.76
	ndn	5.83
	ddn	6.05
	dnd	6.07
	ndd	6.29
ddd	6.33	

Figure 1. Mean responses to the 27 items (conditions) of the EESQ-M.

Note. Items are rank-ordered by their mean rating. The items are labeled by Type of information (in order, Normative, Ipsative, and Expectation) and direction of comparison (d = downward, u = upward, and n = neutral). For example, the item labeled 'und' contains upward normative, neutral ipsative, and downward expectation information.

Table 1. Results of the 3 × 3 × 3 (Normative × Ipsative × Expectation) Multivariate Analysis of Variance (MANOVA).

Effect	Wilks Lambda	F _(df1, df2)	p	eta-squared
Normative	.286	94.8 _(2, 76)	< .001	.714
Ipsative	.138	236.5 _(2, 76)	< .001	.862
Expectation	.122	272.5 _(2, 76)	< .001	.878
Normative × Ipsative	.481	20.0 _(4, 74)	< .001	.519
Normative × Expectation	.363	32.5 _(4, 74)	< .001	.637
Ipsative × Expectation	.275	48.7 _(4, 74)	< .001	.725
Normative × Ipsative × Expectation	.228	29.6 _(8, 70)	< .001	.772

that musicians' performance evaluations are impacted by all three evaluative standards both as main effects and in combination (interaction effects) with each other. However, consistent with our primary hypothesis and Bartók's quotation, the normative standard has a smaller effect size (eta-squared = .714) compared to the ipsative (eta-squared = .862) and expectation (eta-squared = .878) standards. Consistent with this finding is the direct comparison of the three standard influence scores derived from these data as the difference between the mean of the nine downward and the nine upward items for each standard, experimentally controlling for the other

Table 2. Means and standard deviations of, and correlations among, the normative, ipsative, and expectation influence scores.

Score	Mean	SD	Correlations		
			Normative	Ipsative	Expectation
Normative influence	1.66	1.06	–		
Ipsative influence	1.90	.77	.247	–	
Expectation influence	2.09	.79	.227	.640	–

standards. The means and standard deviations of these influence scores, as well as the correlations among them are shown in Table 2. Comparisons of these means indicates that the influence of the expectation standard is significantly greater than the influence of both the normative ($t_{(77)} = 3.24, p = .002$) and ipsative ($t_{(77)} = 2.41, p = .018$) standards on performance evaluation. The ipsative standard is also more influential than the normative standard, although not reaching conventional levels of statistical significance ($t_{(77)} = 1.90, p = .062$). Although the expectation standard is on average the most influential, it is important to note that averages do not apply to all musicians. If we simply count the number of musicians whose influence score was the highest across the three standards, 17 (21.8%) of the musicians were most strongly influenced by the normative standard, 23 (29.5%) were most strongly influenced by the ipsative standard, and 31 (39.7%) were most strongly influenced by the expectation standard. Seven (9%) of the musicians could not be categorized as their highest influence scores were the same for two standards.

The downward and upward scales are the average of the seven items that contain only downward (positive), and the seven items that contain only upward (negative) information, respectively. Not surprisingly, and as shown in Figure 1, musicians evaluated performances that received only downward evaluative information much more positively (Mean = 5.99, $SD = .66$) than performances that received only upward evaluative information (Mean = 2.32, $SD = .81$) ($t_{(77)} = 25.61, p < .001$). However, there also appears to be an asymmetry in the impact of upward and downward information on musician's evaluations. Specifically, the mean downward rating was 1.01 ($SD = .66$) from the top of the scale (7) whereas the mean upward rating was 1.32 ($SD = .81$) from the bottom of the scale (1), and this asymmetry is statistically significant ($t_{(77)} = 3.65, p < .001$). Musicians on average appear to be slightly more responsive to positive evaluative information than negative evaluative information. However, we again recognize that an average tendency does not mean that all musicians are more influenced by positive than negative evaluative information. Of the 78 musicians in our sample, 51 (65%) were influenced more by positive than negative evaluative information. However, 20 (26%) of the musicians showed the opposite tendency to be more influenced by negative information, and 7 (9%) were equally influenced by positive and negative information.

Discussion

The results of this study provide support for both hypotheses. In general, musicians are not as influenced by normative standards compared to ipsative and expectation standards. This result is striking as it stands in contrast to previous work on how students evaluate general academic performance where the normative standard is dominant. However, the quotation from Bartók at the beginning of this article is not entirely correct; musicians are somewhat like horses, as some musicians are still influenced by social comparisons. In our sample, 21.8% of the musicians were found to be most influenced by the competitive standard of social comparison.

Nonetheless, across all of the analyses it was the expectation standard that had the largest average impact on musicians' performance evaluations and that was the dominant standard for the largest number of musicians (40%). Consistent with the second hypothesis, musicians also tended to be more sensitive to positive evaluative information than negative information. What, then, are the implications of these findings?

Social Comparison Theory (Festinger, 1954) has been the dominant theory of how people evaluate their experiences, behaviors, and performances for 60 years. This longevity is remarkable, as according to Hebb, typical half-life for a psychological theory is 5 years (Brown & Milner, 2003). Social comparison, as the basis for self-evaluation, is so ingrained in the psychological literature that the possibility of other standards is rarely considered. Indeed, Gilbert, Giesler, and Morris (1995) showed that in many contexts social comparison occurs spontaneously and automatically. The results of the present research, although clearly indicating that musicians are influenced by normative comparisons, provide a clear counter-example to social comparison. In the case of evaluating musical performances, musicians generally base their evaluations on their expectations rather than on the performances of other musicians. We speculate that musical performances as other types of artistic expression are generally motivated by a person's expectation of what an artist should be. The importance of this result is that it suggests that understanding self-evaluations may require the consideration of other standards besides the normative one. The broadening of Festinger's (1954) social comparison theory into a more general theory of evaluation that includes ipsative and expectation standards seems appropriate as we now have one empirical example where social comparisons are not the dominant basis for self-evaluation. Broadening our approach to self-evaluation is also consistent with recent work by Langer et al. (2010), who specifically discussed the 'mindlessness' of social comparison. They report research that showed that engaging in social comparisons, whether upward or downward, had negative impacts on the creativity of individuals and resulted in less positive evaluations of the participants' creative work (Langer et al., 2010).

In this regard, it is also interesting to consider the 'Lake Wobegon effect' (Kruger, 1999). Lake Wobegon is a fictional town that was used by the comedian Garrison Keillor in his radio show, the *Prairie Home Companion*. Among its characteristics, Lake Wobegon is described as a town where 'all the children are above average.' The idea that all children are 'above average' amuses people with an understanding of statistics because a distribution where all the scores are above the mean is impossible. However, such a distribution is impossible only when a normative standard of comparison is used. If a distribution is based on expectations or self-comparison it is quite possible for everyone's performance and ability to be higher than their average expectation or higher than their average past performance. By expanding the basis of our self-evaluations beyond the normative standard of social comparisons we allow people to recognize positive aspects of a performance that is 'below average' only in a normative sense.

The music industry is a highly competitive one. Despite the competitiveness of the field, our results suggest that the majority of musicians are able to de-emphasize competition and focus instead on the more internal standards of ipsative and especially expectation information for evaluating their performances. The mindlessness of social comparison and the positive aspects of the Lake Wobegon effect, where, using an ipsative or expectation standard all musicians can be above average, suggest that the ability of musicians to de-emphasize the competitive aspects of their field in evaluating their performances is likely adaptive. Alternatively, our results may suggest the difficulty in creating a normative standard for musical performances, because the considerable diversity makes direct comparisons to other performers difficult. Coimbra et al. (2001) noted that the internal schemes that musicians use to grade their performances do not seem to contain any reference to other musician's performances.

The finding that musicians generally give more weight to positive (downward) as opposed to negative (upward) also represents a resiliency in the face of failure that is important for musicians' ultimate success (Knudsen, 2013). Of course, the field of music is no different than nearly all other fields where the ability to overcome failure and criticism is the apparent key to success. There are classic stories of the great guitarist Jimi Hendrix being booed off the stage when he opened for the band The Monkees, of Madonna's first band, The Breakfast Club, being dropped from its record label, and of Elvis Presley being told to keep his job as a truck driver when he appeared on the radio show *Grand Ole Opry*. Although to our knowledge the term has never been used, our results and these stories suggest that there may be an 'adaptive evaluational style' that is similar to the widely studied adaptive attributional style (Anderson, 1999; Peterson et al., 1982) or self-serving bias that characterizes people's tendencies to explain positive events as the result of their own actions and abilities and negative events as the result of factors external to themselves. A hypothesis for future research may be the tendency for people who offer adaptive explanations for events to also exhibit adaptive evaluations for events.

All research has limitations, and the obvious limitations of this research are that it is based on a relatively small sample of college students who were also performing musicians. One direction for future research is replicating these findings on a larger and perhaps more diverse sample of musicians, including those who are experienced professionals. A larger sample might also permit separating the group into different instrument and genre categories. However, despite a relatively small, homogenous sample, we obtained clear and strong support for both of our hypotheses.

Another limitation of this research is that we did not obtain extensive secondary measures that might be correlates of the individual differences we found in this research. Although our primary goals were to evaluate general self-evaluative tendencies of musical performers, these general tendencies occur against a backdrop of individual differences. Most musicians tend to emphasize expectation or ipsative standards, but we are very interested in the minority of the musicians that emphasize the more competitive standard of social comparison. Are these musicians the performers who tend to win competitions or get the best grades? Do they enjoy performing more or less? Are they more or less motivated to continue in professional careers? The same questions may be asked of the minority of musicians who exhibit the 'less adaptive' evaluational style of emphasizing negative feedback? Future research can begin to clarify the causes and consequences of individual differences in evaluational styles.

Conclusion

The study reported here was based on an experimental design that allowed us to disentangle the influence of three different evaluative standards based on normative, ipsative, and expectation information. To our knowledge, the results of this study are among the first where the normative standard of social comparison did not prove to have the strongest influence on people's evaluations. Of course, we selected the domain of musical performance because the intuitions of musicians suggested that competitive standards were not common in how musicians evaluate their performances. Based on our data, the intuitions of musicians are supported. Using non-competitive standards such as self-improvement and expectations likely has adaptive consequences for self-esteem, motivation, and quality of life as such standards truly allow all musicians, just as the children in Lake Wobegon, to be above average.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

- Albert, S. (1977). Temporal comparison theory. *Psychological Review*, 84, 425–503.
- Anderson, C. A. (1999). Attributional style, depression, and loneliness: A cross-cultural comparison of American and Chinese students. *Personality and Social Psychology Bulletin*, 25, 482–499.
- Aspinwall, L. G., & Taylor, S. E. (1993). Effects of social comparison direction, threat, and self-esteem on affect, self-evaluation, and expected success. *Journal of Personality & Social Psychology*, 64, 708–722.
- Béla Bartók. (n.d.). BrainyQuote.com. Retrieved from <http://www.brainyquote.com/quotes/quotes/b/BélaBartók190412.html>
- Brown, R. E., & Milner, P. M. (2003). The legacy of Donald O. Hebb: More than the Hebb synapse. *Nature Reviews: Neuroscience*, 4, 1013–1019.
- Chaplin, W. F., & Buckner, K. (1988). Self-ratings of personality: A naturalistic comparison of normative, ipsative, and idiothetic standards. *Journal of Personality*, 56(3), 509–530.
- Coimbra, D., Davidson, J., & Kokotsaki, D. (2001). Investigating the assessment of singers in a music college setting: The students' perspective. *Research Studies in Music Education*, 16(1), 15–32.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140.
- Gilbert, D., Giesler, R., & Morris, K. (1995). When comparisons arise. *Journal of Personality and Social Psychology*, 69(2), 227–236.
- Goolsby, L., & Chaplin, W. F. (1988). The impact of normative, ipsative, and idiothetic information on feelings about academic performance. *Journal of Research in Personality*, 22, 445–464.
- Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, 94(3), 319–340.
- Knudson, J. (2013). *Successful Failures: Wisdom for Aspiring Musicians*. Northern Ireland: David James Publishing.
- Kruger, J. (1999). Lake Wobegon be gone! The 'below-average effect' and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology*, 77, 221–232.
- Lamiell, J. T. (1981). Toward an idiothetic psychology of personality. *American Psychologist*, 36, 276–289.
- Langer, J., Delizonna, L., & Pirson, M. (2010). The mindlessness of social comparisons and its effect on creativity. *Psychology of Aesthetics, Creativity, and the Arts*, 4, 68–74.
- Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, 41, 954–969.
- McArthur, L. A. (1972). The how and what of why: Some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology*, 22, 171–193.
- McCormick, J., & McPherson, G. (2003). The role of self-efficacy in a musical performance examination: An exploratory structural equation. *Psychology of Music*, 31, 37–52.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231–259.
- Peterson, C., Semmel, A., Von Bayer, C., Abramson, L., Metalsky, G., & Seligman, M. E. P. (1982). The attributional style questionnaire. *Cognitive Therapy and Research*, 6(3), 287–300.
- Priest, T. (2006). Self-evaluation, creativity, and musical achievement. *Psychology of Music*, 34(1), 47–61.
- Ritchie, L., & Williamon, A. (2011). Measuring distinct types of musical self-efficacy. *Psychology of Music*, 39(3), 328–344.
- Suls, J., & Wheeler, L. (2000). *Handbook of social comparison*. New York, NY: Kluwer Academic.
- Suls, J. E., & Wills, T. A. E. (1991). *Social comparison: Contemporary theory and research*. Hillsdale, NJ: Erlbaum.
- Taylor, S. E., Wayment, H. A., & Carrillo, M. (1996). Social comparison, self-regulation, and motivation. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition, Vol. 3: The interpersonal context* (pp. 3–27). New York, NY: Guilford Press.
- Thompson, S., & Williamon, A. (2003). Evaluating evaluation: Musical performance assessment as a research tool. *Music Perception: An Interdisciplinary Journal*, 21(1), 21–41.
- Zajonc, R.B. (1980). Feeling and thinking preferences need no inferences. *American Psychologist*, 35(2), 151–175.

Appendix. EESQ-M

Evaluating your musical performance

In this study we are interested in how you evaluate your musical performances. On this questionnaire you will be given 27 items that contain different information that you might use to evaluate the following event.

You gave a musical performance.

Although each of the items will concern this event, each item will contain different information about the event **and you should treat each item separately. Also, you will not be told about how your performance was evaluated by critics or teachers, the difficulty of the piece you performed, or what type of music you performed and we do not want you to assume anything specific about the performance.**

Instead, the information we provide about the performance will be based on some relative standards. Specifically you will be told how well or poorly you performed: 1) relative to other people’s performances, 2) relative to how you expected or wanted you performance to go and/or 3) relative to your past performances. Thus, each item will contain either one, two, or three types of relative information about your performance. Please think about **all** of the information provided in an item before responding to that item.

Read each of the following items carefully and take a moment to imagine yourself performing. Then, think about the relative standards we ask you to use in each item to evaluate the performance. Finally, on the basis of those standards, rate how you would feel about your performance circling a number on the following scale:

I would feel:

1	2	3	4	5	6	7
Extremely Bad	Very Bad	Somewhat Bad	Neutral	Somewhat Good	Very Good	Extremely Good

Remember, although they concern the same general event, each item is **independent** of the others. Therefore, your response to each item should be made without any influence from the other items. To emphasize that they are separate, each item is separated by a thick line.

Event: You gave a musical performance.

1.

(No information given)—How do you evaluate performing in general?

I would feel:

1	2	3	4	5	6	7
Extremely Bad	Very Bad	Somewhat Bad	Neutral	Somewhat Good	Very Good	Extremely Good



2. Your performance was much poorer than performances you have given in the past.
3. Your performances was much poorer than other people's performances.
and
Your performance was much better than you expected it to be.
4. Your performance was much better than you expected it to be.
and
Your performance was much better than other people's performances.
and
Your performance was much poorer than performances you have given in the past.
5. Your performance was much better than other people's performances.
and
Your performance was much poorer than you expected it to be.
and
Your performance was much poorer than performances you have given in the past.
6. Your performance was much better than you expected it to be.
and
Your performance was much better than performances you have given in the past.
7. Your performance was much poorer than performances you have given in the past.
and
Your performance was much poorer than you expected it to be.
and
Your performance was much poorer than other people's performances.
8. Your performance was much better than performances you have given in the past.
and
Your performance was much better than other people's performances.
9. Your performance was much poorer than you expected it to be.
and
Your performance was much better than performances you have given in the past.
and
Your performance was much poorer than other people's performances.
10. Your performance was much better than other people's performances.
and
Your performance was much poorer than you expected it to be.
11. Your performance was much better than performances you have given in the past.
and
Your performance was much poorer than other people's performances.
12. Your performance was much better than performances you have given in the past.
13. Your performance was much better than you expected it to be.
and
Your performance was much poorer than other people's performances.
and
Your performance was much poorer than performances you have given in the past.

14. Your performance was much better than other people's performances.
15. Your performance was much poorer than you expected it to be.
and
Your performance was much poorer than performances you have given in the past.
16. Your performance was much poorer than performances you have given in the past.
and
Your performance was much better than you expected it to be.
17. Your performance was much poorer than you expected it to be.
and
Your performance was much better than other people's performances.
and
Your performance was much better than performances you have given in the past.
18. Your performance was much poorer than you expected it to be.
and
Your performance was much poorer than other people's performances.
19. Your performance was much better than performances you have given in the past.
and
Your performance was much poorer than you expected it to be.
20. Your performance was much poorer than other people's performances.
and
Your performance was much poorer than performances you have given in the past.
21. Your performance was much poorer than you expected it to be.
22. Your performance was much poorer than performances you have given in the past.
and
Your performance was much better than other people's performances.
23. Your performance was much poorer than other people's performances.
24. Your performance was much better than other people's performances.
and
Your performance was much better than you expected it to be.
and
Your performance was much better than performances you have given in the past.
25. Your performance was much better than you expected it to be.
and
Your performance was much poorer than other people's performances.
and
Your performance was much better than performances you have given in the past.
26. Your performance was much better than you expected it to be.
27. Your performance was much better than other people's performances.
and
Your performance was much better than you expected it to be.