Self-Reported GPA and SAT: A Methodological Note

Jerrell C. Cassady

Recommended Citation

Cassady, Jerrell C. (2000) "Self-Reported GPA and SAT: A Methodological Note," *Practical Assessment, Research, and Evaluation*: Vol. 7 , Article 12.
DOI: https://doi.org/10.7275/5hym-y754
Available at: https://scholarworks.umass.edu/pare/vol7/iss1/12

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Self-Reported GPA and SAT: A Methodological Note

Jerrell C. Cassady
Ball State University

The use of self-reports from students is a common, yet risky, methodological venture in psychological research. Relying upon an individual to provide an accurate and unbiased rating of her or his ability, attitude, or past experiences is problematic, considering that the behavioral revolution in psychology was driven in part by a rejection of the reliance on malleable data sources. Researchers now concentrate on observable phenomena that can be validated; however, they are often forced to ask participants to provide various forms of data without having the means to verify their accuracy.

The research base demonstrating the magnitude of disparity between actual and reported performance scores is very limited. This article investigates the methodological practice of relying on self-reported Scholastic Assessment Test (SAT) and college grade point average (GPA) scores provided by undergraduate research participants. It also attempts to provide an explanation for the differential reliability of self-reported SAT and GPA values by examining the differences in students' access to these two scores. Because most undergraduates see their GPAs every 16 to 20 weeks, but seldom review their SAT scores after college entrance, it is expected that reliability will be greater for self-reported GPA scores than for self-reported SAT scores among undergraduates.

The research in SAT score accuracy has generally indicated that students' reports correlate with actual scores in the range of .60 to .80 (Goldman, Flake, and Matheson, 1990; Frucot and Cook, 1994; Trice, 1990). Furthermore, there is evidence that individuals who do not provide their scores are more likely to have low SAT scores, suggesting a potential skew in the self-report performance literature (Flake and Goldman, 1991; Trice, 1990). In a rigorous analysis of the relationship between actual and reported scores on the SAT, Shepperd (1993) reported that students with low SAT scores not only inflated their self-reported scores, but also rated the score they received on the SAT as inaccurate or flawed. Furthermore, when students reported SAT scores with no explicit instructions, the tendency to inflate the score was evident. However, when the students were asked to report their SAT scores for a second time (two months after the initial report), but with an incentive for accuracy and the assurance that any inflation would be known, the average deviation from true score was 9 points for the total scale SAT score, a mere 1/10 of a standard deviation (Shepperd, 1993). Shepperd hypothesized that this pattern supported the theory that the inflation was an attempt to portray a positive image, rather than a misrepresentation arising from a memory deficit. As for GPA ratings, there is also evidence for skewed self-reports; specifically, there is greater inflation by students with lower GPAs than by students with higher GPAs (Dobbins, Farh, and Werbel, 1993; Frucot and Cook, 1994). This inflation of GPA has been found to be free from a ceiling effect, and it has been proposed to be a consequence of social desirability (Dobbins et al., 1993).

Altogether, there are insufficient data regarding the validity of self-reported SAT and GPA values to be confident in this common methodological practice. Furthermore, the existing reports often do not provide an indication of the absolute, or relative, degree of deviation. This study was conducted to test the accuracy and trends of deviation noted in undergraduates' self-reported SAT and GPA values. The magnitude of deviation was examined through two independent variables, the direction of deviation (if any) and the actual performance level of the student on the measure in question. The results were expected to support previous reports that self-reported values for GPA and SAT were relatively reliable (in the range of .70 to .90). Furthermore, the results were expected to identify that for both GPA and SAT, low scorers' ratings would vary from actual scores more than high scorers, with the self-reported values demonstrating an inflated value. Finally, it was predicted that individuals who overestimated their performance levels would do so to a greater magnitude than those individuals reporting an underestimation of their performance.

Method

Participants

Eighty-nine undergraduate students at a mid-sized Midwestern university reported their current cumulative GPA and the scores they had received on the SAT. Ninety-six percent of the participants ($\underline{n}$ = 86) identified themselves as Caucasian; the remaining three students identified themselves as African American. Eighty-nine percent ($\underline{n}$ = 79) were female. All participants were in the second year of the undergraduate preservice teacher education program. Participants reported ages ranging from 19 to 28 ($\underline{M}$ = 19.99, $\underline{SD}$ = 1.06).

Procedure

The participants were asked to provide their undergraduate cumulative GPAs and their official SAT scores as part of another research project. Those participants who indicated they were unsure of their scores were asked to provide their "best guess" regarding the SAT verbal, math, and total scores, as well as GPA. There was no indication to the students at the time they provided their scores what the

scores would be used for, or that the scores would be checked against their official records. The participants were debriefed in a subsequent experimental session, at which time they provided consent to access the necessary university records.

Students who did not take the SAT (typically taking the ACT for entrance) were excluded from the analyses of SAT score accuracy. Similarly, students without official university grade records (i.e., transfer students from community colleges) were excluded from the analyses on the accuracy of GPA. After the participants granted consent to access the records, their actual SAT and GPA values were gathered from official university records during the same semester to ensure that the students' official GPA was not affected by grades not finalized at the time of data collection.

Analyses

Initial investigation of the data relied on examining the correlation between the self-reported scores and official records to establish an overall level of accuracy. In addition, the analyses targeted two additional potential group differences, tendency to deviate and actual performance levels.

To investigate the impact of direction of reported scores' deviation from the actual scores, each participant's reported value was categorized as either an overestimation, an underestimation, or as accurate. These reports were examined to identify whether the magnitude of deviation from students who overestimated and underestimated their scores differed significantly from each other.

The second grouping variable was based on actual performance levels. To examine whether low-scoring individuals inflated their scores more than high-scoring individuals in both SAT and GPA self-report values, four groups were established for each measure, using the quartile split method.

To investigate differential magnitudes of deviation based on both direction of deviation (overestimation and underestimation) and actual performance level, univariate analyses of variance were conducted on the absolute value of the deviation of the reported score from the actual score. The use of absolute values is appropriate because the direction of deviation is represented in the directional grouping variable.

Results

Students' self-reported GPA scores were found to be remarkably similar to official records. The Pearson product moment correlation revealed a significant correlation between self-reported and actual cumulative GPA, $r = .97$, $p < .0001$, $n = 75$. Similarly, correlational analyses of the accuracy of the students' self-reported SAT scores revealed significant relationships between self-reports and actual performance levels for the total score ($r = .88$, $p = .0001$, $n = 72$), verbal subscale ($r = .73$, $p = .0001$, $n = 64$), and math subscale ($r = .89$, $p = .0001$, $n = 64$).

To examine deviation of GPA scores, a two by four univariate analysis of variance was used, with two levels of direction of deviation (overestimation and underestimation) and four levels of actual GPA (as established by quartile placement in the sample). The ANOVA revealed a significant main effect for level of GPA on deviation from reality. Neither the main effect for direction of deviation nor the interaction produced a significant effect (see Table 1). The data indicated progressively more accurate ratings of GPA as the level of GPA increased (see Table 2 for means and standard deviations). Post-hoc analyses of group differences revealed differences between the quartiles, with the first quartile deviations being significantly higher than the third ($p < .005$) and fourth ($p < .001$), and the participants in the second quartile producing significantly higher deviations than the fourth ($p < .05$).

Table 1: Analysis of variance for group differences in deviation of self-reported GPA

Source	df	F
Direction of Deviation (D)[a]	1	.49
Performance Level (L) [b]	3	5.06***
D X L	3	1.38
Error	60	(.01)

Note: The figure in parentheses represents the mean square error. *** - $p < .001$, ** - $p < .01$, *- $p < .05$. [a] Direction of deviation includes overestimation and underestimation. [b] Performance level was determined by quartile splits on GPA.

Table 2: Deviations of self-reported scores from actual scores for GPA and SAT subscale scores

	GPA[a]			SAT Verbal			SAT Math		
Group	n	M	SD	n	M	SD	n	M	SD
Underestimation									
First Quartile	10	.117	.134	3	53.33	75.06	5	40.00	31.62
Second Quartile	8	.120	.118	4	95.00	137.23	3	13.33	5.77
Third Quartile	15	.064	.057	8	31.25	28.00	5	18.00	4.47

Fourth Quartile	11	.050	.038	4	22.50	25.00	3	40.00	20.00

Overestimation

First Quartile	8	.221	.184	9	70.00	41.83	6	65.00	37.28
Second Quartile	9	.126	.076	4	37.50	15.00	6	36.67	20.66
Third Quartile	5	.065	.027	5	31.00	14.32	4	22.50	15.00
Fourth Quartile	4	.023	.020	3	80.00	81.85	5	36.00	19.49

Note: Reported scores are the absolute value of the difference between the actual and reported values. [a] GPA score was based on 4.0 scale.

Similar analyses were conducted on the verbal and math subscales of the SAT. Because the total score for the SAT is a combination of these two subscales, no additional analysis of the total score was conducted. To examine deviation of the SAT subscale scores, two separate two by four univariate analyses of variance were conducted, with two levels of direction and four levels of SAT performance. The ANOVA revealed no significant effects for the verbal subscale (see Table 3). The results for the math subscale revealed a trend similar to GPA, with a significant main effect for level of SAT performance (as determined by quartile placements), while the main effect for direction of deviation and interaction were not significant (see Table 3). Post-hoc analyses revealed that members in the first quartile produced significantly higher deviations than members in the second ($p < .03$) and third ($p < .004$) quartiles.

Table 3: Analyses of variance for group differences in deviation of self-reported SAT subscales

Source	df	F
SAT Verbal		
Direction of Deviation (D)[a]	1	.14
Performance Level (L) [b]	3	1.10
D X L	3	.79
Error	31	(2963.68)
SAT Math		
Direction of Deviation (D)[a]	1	2.30
Performance Level (L) [b]	3	3.66*
D X L	3	.78
Error	29	(559.48)

Note: The figures in parentheses represent the mean square error. *** - $p < .001$, ** - $p < .01$, *- $p < .05$.
[a] Direction of deviation includes overestimation and underestimation [b] Performance level was determined by quartile splits on the SAT subscales.

Discussion

The results of this study of GPA and SAT self-reports allow for a general statement regarding the role of self-reported performance indicators to be made. The initial hypothesis regarding accuracy of ratings was supported, revealing that the participants had highly reliable ratings of cumulative GPA ($r = .97$). Such high correlations would suggest that overall, self-reported GPA levels are sufficiently accurate. The overall accuracies of the students' self-reported SAT scores were considerably lower than the accuracy of GPA; however, the average accuracy was still within reasonable guidelines (Nunnally and Bernstein, 1994). The results supported the expectation that the accuracy of self-reported SAT scores would be lower than self-reported GPAs.

This difference in accuracy may be related to the factors of repetition and recency. Cumulative GPA is reported to undergraduate students on a consistent and frequent basis, typically at least two to three times per year. SAT scores, however, are not typically reported to the students once they've been admitted to the university; consequently, the majority of these participants would not likely have seen their official SAT scores for a period of two or more years.

Further investigation revealed that accuracy of self-reported scores was dependent upon the independent variable of performance level. The analyses of accuracy in self-reported GPA revealed that the bottom 25% of students provided estimates that were significantly less accurate than each of the remaining quartile groups. These data support a trend reported by Dobbins et al. (1993), who revealed that students with lower GPAs tended to inflate their scores more than students with higher averages. In a similar vein, self-reports of SAT performance generally became more accurate as actual performance increased. Overall, it appears that students at the lowest end of performance are more likely than the high-achieving groups to misrepresent their scores. This is consistent with the proposal that the students at the low-performing levels may provide inflated scores as a function of social desirability (Dobbins et al, 1993).

Contrary to the initial hypothesis, there were no differences in deviation from actual scores by those participants who overestimated and underestimated their performance levels. The expectation was that the deviations would be higher for overestimators, consistent with the social desirability hypothesis. However, no such trend was revealed, suggesting that the deviations from actual scores are due in part to errors in memory, and not all deviations are driven by a desire to misrepresent ability levels.

Given ideal conditions, there would be no sense in relying on students to report their GPA and SAT scores from memory. However, several conditions may limit a researcher's ability to gain access to official records, including administrative rules and privacy issues. When these conditions arise, forcing a researcher into a compromised methodological activity, these data suggest that researchers can rely upon self-reported GPA estimates. The data suggest that the use of self-reported SAT scores is less reliable than GPA estimations, but can be tempered by indicating to the students that accuracy is of primary interest, perhaps by assuring anonymity to the participants (see Shepperd, 1993). The use of self-reported GPA and SAT scores increases the efficiency of data collection available to researchers, particularly when these scores are simply additional variables of interest, perhaps when attempting to account for variance in designs examining course performance, test anxiety, or career orientations. The ease of acquiring these values through self-report, combined with the high levels of accuracy under the current methodology, make this practice an enticing alternative to the more laborious process of accessing official student records.

However, these results do not support the use of self-reported GPA and SAT scores for policy decisions, particularly if the students are able to determine the intent of the score collection. In situations where the students' GPA and SAT scores will be used to differentiate among candidates for selection into special programs or positions, students may be more likely to provide false estimates to improve their standing. Furthermore, this practice should not be generalized to participants at different developmental levels without assessing a pilot sample to ensure the reliability is still adequate.

References

Dobbins, G. H., Farh, J. L., and Werbel, J. D. (1993). The influence of self-monitoring and inflation of grade-point averages for research and selection purposes. *Journal of Applied Social Psychology, 23,* 321-334.

Flake, W. L., and Goldman, B. A. (1991). Comparison of grade point averages and SAT scores between reporting and nonreporting men and women and freshmen and sophomores. *Perceptual and Motor Skills, 72,* 177-178.

Frucot, V. G., and Cook, G. L. (1994). Further research on the accuracy of students' self-reported grade point averages, SAT scores, and course grades. *Perceptual and Motor Skills, 79,* 743-746.

Goldman, B. A., Flake, W. L., and Matheson, M. B. (1990). Accuracy of college students' perceptions of their SAT scores and high school and college grade point averages relative to their ability. *Perceptual and Motor Skills, 70,* 514.

Nunnally, J. C., and Bernstein, I. H. (1994). *Psychometric Theory* (3rd Ed.). New York: McGraw-Hill, Inc.

Shepperd, J. A. (1993). Student derogation of the Scholastic Aptitude Test: Biases in perceptions and presentations of College Board scores. *Basic and Applied Social Psychology, 14,* 455-473.

Trice, A. D. (1990). Reliability of students' self-reports of scholastic aptitude scores: Data from juniors and seniors. *Perceptual and Motor Skills, 71*, 290.

Address all correspondence to Jerrell C. Cassady, Ph.D., Department of Educational Psychology, Ball State University, Muncie, IN 47306; jccassady@bsu.edu