

# Conflict of interest explains the size of student evaluation of teaching and learning correlations in multisection studies: A meta-analysis

Bob Uttl<sup>Corresp., 1</sup>, Kelsey Cnudde<sup>2</sup>, Carmela A White<sup>3</sup>

<sup>1</sup> Psychology, Mount Royal University, Calgary, AB, Canada

<sup>2</sup> Psychology, University of Calgary, Calgary, AB, Canada

<sup>3</sup> Psychology, University of British Columbia Okanagan, Kelowna, BC, Canada

Corresponding Author: Bob Uttl  
Email address: uttlbob@gmail.com

We examined the associations between the size of student evaluation of teaching and learning (SET/learning) correlations reported in multisection studies as reviewed by Uttl, White, and Wong Gonzalez (2017) and publication period (prior to 1981 vs. 1981 and after) and presence of several conflicts of interest (COIs) including corporate, administrative, evaluation unit, SET author, and funder interests. Our meta-analyses of SET/learning correlations reported by multisection studies show that researchers with a vested interest in finding large positive SET/learning correlations found, on average, large positive SET/learning correlations. In contrast, researchers with no identifiable COIs found that SET/learning correlations were zero or nearly zero. The largest SET/learning correlations were reported by authors with ties to SET selling corporations. Smaller but still substantial SET/learning correlations were reported by researchers with administrative assignments and researchers with evaluation units/departments responsible for the administration of SETs. Moreover, authors with the most significant COIs were publishing their studies primarily prior to 1981 whereas authors with no or less significant COIs were publishing their studies in 1981 or afterwards. Studies published prior to 1981 reported small but significant ( $r = .31$ ) SET/learning correlations whereas studies published in 1981 and after reported near zero, non-significant SET/learning correlations ( $r = .06$ ). The presence of COIs was associated with earlier publication date but also with smaller samples. Finally, whereas corporate, administrative, and evaluation unit authors nearly ceased publishing multisection studies on SET/learning correlations, authors from business and economics departments are now responsible for the substantial portion of newer, larger, and higher quality studies published in 1981 and after.

Conflict of interest explains the size of student evaluation of teaching and learning correlations in  
multisection studies: A meta-analysis

Bob Uttl<sup>1</sup>, Kelsey Cnudde<sup>2</sup>, & Carmela A. White<sup>3</sup>

<sup>1</sup>Mount Royal University, Calgary, AB, Canada

<sup>2</sup>University of Calgary, Calgary, AB, Canada

<sup>3</sup>University of British Columbia Okanagan, Kelowna, BC, Canada

Corresponding Author:

Bob Uttl<sup>1</sup>

Email address: [uttlbob@gmail.com](mailto:uttlbob@gmail.com)

# Abstract

We examined the associations between the size of student evaluation of teaching and learning (SET/learning) correlations reported in multisection studies as reviewed by Uttl, White, and Wong Gonzalez (2017) and publication period (prior to 1981 vs. 1981 and after) and presence of several conflicts of interest (COIs) including corporate, administrative, evaluation unit, SET author, and funder interests. Our meta-analyses of SET/learning correlations reported by multisection studies show that researchers with a vested interest in finding large positive SET/learning correlations found, on average, large positive SET/learning correlations. In contrast, researchers with no identifiable COIs found that SET/learning correlations were zero or nearly zero. The largest SET/learning correlations were reported by authors with ties to SET selling corporations. Smaller but still substantial SET/learning correlations were reported by researchers with administrative assignments and researchers with evaluation units/departments responsible for the administration of SETs. Moreover, authors with the most significant COIs were publishing their studies primarily prior to 1981 whereas authors with no or less significant COIs were publishing their studies in 1981 or afterwards. Studies published prior to 1981 reported small but significant ( $r = .31$ ) SET/learning correlations whereas studies published in 1981 and after reported near zero, non-significant SET/learning correlations ( $r = .06$ ). The presence of COIs was associated with earlier publication date but also with smaller samples. Finally, whereas corporate, administrative, and evaluation unit authors nearly ceased publishing multisection studies on SET/learning correlations, authors from business and economics departments are now responsible for the substantial portion of newer, larger, and higher quality studies published in 1981 and after.

*It is difficult to get a man to understand something when his salary depends on his not understanding it.* (Upton Sinclair)

*Money makes the world go around, the world go around, the world go around, money makes the world go around, it makes the world go 'round. A mark, a yen, a buck, or a pound, a buck or a pound, is all that makes the world go around, that clinking clanking sound, can make the world go 'round.* (John Kander, from "Cabaret")

For decades, colleges and universities have been asking students to evaluate teaching effectiveness of their professors using a variety of Student Evaluation of Teaching (SET) questionnaires based on a widespread belief that students learn more from more highly rated professors. The SETs are typically administered during the last few weeks of classes, before students receive their final grades, and ask students to rate their professors' teaching, for example, their overall teaching effectiveness as well as their knowledge, clarity, organization, friendliness, fairness, approachability, availability, etc. The SET ratings for each course/class are summarized, often by calculating means for each rated item and across all items, and the mean SETs are then used as a measure of professors' teaching effectiveness. The SET ratings are subsequently used in a variety of high stakes personnel decisions such as faculty hiring, firing, promotion, tenure, merit pay, teaching awards, and distinguished professor awards (Spooren, Brockx, & Mortelmans, 2013; Stark & Freishtat, 2014; Uttl et al., 2017; Wachtel, 1998).

The main evidence cited for the belief that students learn more from professors with higher SET ratings were several meta-analyses of multisection studies showing small-to-moderate correlations between SET ratings and student learning (Uttl et al., 2017). For example, Cohen (1981) claimed that the correlation between overall instructor SET ratings and learning was .43 to .44. Similarly, Feldman (1989) claimed that the correlations between various SET item ratings and learning were as high as .57. These meta-analyses were reviewed by numerous researchers who all concluded that these meta-analyses are strong evidence of the validity of SET ratings as a measure of student learning. For example, Abrami and d'Apollonia (1999, p. 519) wrote: "[SET] ratings explain instructor impact on student learning to a moderate extent (corrected  $r = .47$ )"; Marsh (Marsh, 2007, p. 339) wrote: "Perhaps more than any other area of SET research, results based on the multisection validity paradigm support the validity of SET."; Benton and Cashin (2012, p. 4) wrote: "the multisection studies show that classes in which the students gave the instructor higher ratings tended to be the ones where the students learned more". Not surprisingly, these summary reviews then made their way as established facts into various self-help books for starting faculty members. For example, Davis (2009, p. 534) wrote: "Students of highly rated teachers achieve higher final exam scores, can better apply course material, and are more inclined to pursue the subject subsequently."

Recently, Uttl and colleagues (2017) published a detailed review of all previously published meta-analyses of the SET/learning correlations and found that they all suffered from multiple critical methodological shortcomings. For example, none of these meta-analyses adequately considered the possibility that their results may be an artifact of small sample bias

even though the smallest studies (with 5 sections) showed correlations as high as .91 (“voodoo”) whereas the larger studies (with 30 or more sections) showed much smaller correlations. The substantial relationship between the sample size and the size of the SET/learning correlations would have been obvious to anyone who plotted these graphs (Uttl et al., 2017). Moreover, the previous meta-analyses did not adequately describe the search for, and identify all, primary studies, often did not provide effect sizes and sample sizes for each primary study, often failed to weigh SET/learning correlations by sample size, and in general, were so poorly described as to be unreplicable. In one remarkable review of previous meta-analyses, Abrami, Cohen and d’Apollonia (1988) were concerned about “disparate conclusions of the [multisection study] reviews”, discussed “why the reviews differ”, and even considered a possibility of publication bias (one type of small-sample bias) but dismissed it. They puzzled over “troublesome” disagreement between SET/learning correlations extracted by Cohen (1983) and McCallum (1984). However, instead of checking the extracted values against the primary studies and determining the source of the disagreement, Abrami et al. chose to speculate about possible sources of disagreement. As Uttl et al. (2017) found, the principal reason for the “troublesome” disagreement was McCallum’s (1984) failure to correctly extract relevant values from primary studies.

When Uttl et al. (2017) re-analyzed the previously published meta-analyses of multisection studies, they found that previous findings were artifacts of small sample bias. The small sample size studies showed large SET/learning correlations whereas the large sample size studies showed no or only minimal SET/learning correlations. When small sample bias was taken into account through a variety of meta-analytic methods, the estimated SET/learning correlations were much smaller than previously reported. Moreover, given the shortcomings of the previous meta-analyses, Uttl et al. conducted a new up-to-date meta-analysis of the SET/learning correlations published in multisection studies from the ground up. Their search yielded 51 articles with 97 multisection studies. Using limit meta analysis to adjust for prevalent small sample bias, the new data set resulted in a SET/learning correlation of only  $r = .12$ , 95% CI = (.03, .21). Moreover, when they re-ran the same analyses but only for studies with prior knowledge/ability adjustments, the limit meta-analysis resulted in a SET/learning correlation of  $r = -.06$ , 95% CI = (-.17, .07). Accordingly, Uttl et al. concluded that the multisection studies do not support the claims that students learn more from more highly rated professors. Rather, the multisection studies are evidence that SETs and learning are unrelated.

Why is it that for three to four decades fatally flawed meta-analyses of multisection studies by Cohen (1981) and Feldman (1989) were cited as evidence of SET/learning correlations, and that SETs are valid enough to be used in high stakes personnel decisions including terminating people’s careers? How is it possible that numerous experts in the field who reviewed these meta-analyses did not notice numerous red flags in the previous meta-analyses of multisection studies (e.g., not reporting SET/learning correlations and sample sizes for primary studies, not weighing SET/learning correlations by sample size) as well as in the primary studies themselves (e.g., impossibly high voodoo correlations in small sized primary studies)? How is it possible that when experts noticed “troublesome” disagreement between the SET/learning correlations in different meta-analyses they did not follow up and identify the source of the disagreement and

chose to speculate about it instead?

One possibility, that we investigate in this article, is that the primary findings and the written reviews have been strongly influenced by conflicts of interest (COI). First, even a superficial glance makes it apparent that many multisection studies were authored by researchers working for corporations selling SET systems. For example, John Centra was heavily involved in the development of the Student Instructional Report (SIR) and SIR-II SET systems for Educational Testing Service; Peter Frey developed the Endeavor Instructional Rating Card and founded Endeavor Information Systems, Inc.; Herbert Marsh developed the Student Evaluation of Educational Quality (SEEQ) SET system and founded and became President of Evaluation, Testing and Research, Inc.; and Lawrence Aleamoni developed Aleamoni Course/Instructor Evaluation Questionnaire (CIEQ) and founded and became President of Comprehensive Data Evaluation Services Inc (Hill, 2006; Aleamoni v. Commissioner of Internal Revenue Service). Second, the experts who reviewed the previous meta-analyses of multisection studies and found that the meta-analyses were strong evidence of the validity of SETs as a measure of teaching effectiveness had strong interests in that particular conclusion. For example, Cohen (1981) authored the very first and the most highly cited meta-analysis of multisection studies claiming that SET/learning correlations were greater than  $r = .40$  and immediately thereafter became Assistant Director of the Office of Instructional Services and Educational Research. Benton worked for the IDEA Center, a nonprofit organization selling IDEA SETs to colleges and universities worldwide. Marsh and Aleamoni, as noted above, both developed their own SET systems and founded companies to distribute them. Third, a number of authors of multisection studies worked for units or departments responsible for the evaluation of faculty's teaching effectiveness in various colleges and universities. Finally, some multisection studies were funded by grants from corporations with vested interests, for example, Endeavor Information Systems Inc..

Notably, conflicts of interest tying authors to their SET enterprises are typically not declared and readers may not know about them. For example, we only found out that Aleamoni was the president of Comprehensive Data Evaluation Services Inc. -- a corporation that publishes and distributes Aleamoni's CIEQ SET -- by discovering the case Aleamoni v. Commissioner of Internal Revenue Service where this information is disclosed. He, his wife, and his children owned 100% of shares in this company.

Even though Clayson's (2009) findings are uninterpretable and his conclusions unwarranted (Uttl et al., 2017), Clayson (2009) suggested that the magnitude of SET/ learning correlations may vary by publication year, authors' department, and other factors. Our own review of multisection studies suggests that, in addition to study size and a variety of conflicts of interest, SET/learning correlations may be larger for studies published prior to Cohen's (1981) meta-analysis compared to those published after, as well as studies originating from educational or psychology departments vs. studies originating from business and economics departments. In general, studies published prior to 1981 were conducted primarily by SET corporations and evaluation units, whereas studies published after 1981 were conducted primarily by authors from business and economics departments.

Accordingly, our study had several main goals. The first goal was to determine whether

SET/learning correlations reported in primary studies were larger for studies published prior to 1981 vs. 1981 and after. Cohen's (1981) meta-analysis appeared to have cemented a belief that SET/learning correlations were substantial. However, as noted above, a majority of studies published prior to 1981 were published by authors with corporate, administrative, and evaluation unit COIs, and thus, to the extent to which COIs play a role, we would expect the earlier studies to report larger SET/learning correlations. In addition, earlier vs. later studies employed smaller samples, and thus, in combination with publication bias, earlier studies are likely to report larger SET/learning correlations. The second goal was to examine whether SET/learning correlations reported by authors with the most significant COI -- authors associated with and working for SET selling corporations -- are larger than correlations reported by other authors. Again, to the extent to which COIs explain variability in SET/learning correlations reported by primary studies, we would expect larger SET/learning correlations reported by corporate authors than by other authors. The third goal was to examine whether SET/learning correlations vary with different COIs, including corporate, administrative, evaluation unit, and SET author COIs. The fourth goal was to examine associations between SET/learning correlations and authors' place of employment/work including corporate, administrative, evaluation unit, education and psychology departments, business and economics departments, and other. Our final goal was to see if SET/learning correlations vary with the number of COIs present.

## Method

A preliminary review of the multisection studies revealed that the articles themselves rarely declared conflict of interest and often failed to even disclose sufficient information to determine whether a conflict of interest existed. Accordingly, to determine the presence of any conflict of interest, we proceeded in three steps. First, we examined each article reporting a multisection study. Second, we examined all articles published by the study authors within five years of the multisection study publication date. And third, we searched Google for author's CVs, and other relevant publications.

For each multisection study reported in Uttl et al. (2017), we coded for the presence or absence of any of the following direct and indirect conflicts of interest:

**Corporate interest.** If at least one author worked for a corporation or organization selling SET services (e.g., Educational Testing Service, IDEA), the corporate interest was coded as present.

**Evaluation unit interest.** If at least one author worked for a corporate or university teaching evaluation unit, the evaluation unit interest was coded as present.

**Administrative conflict of interest.** If at least one author was an administrator, for example, a chair of the department or a dean, the administrative interest was coded as present.

**SET author interest.** If at least one author authored or co-authored a widely used SET, the SET author interest was coded as present.

**Funder interest.** If a funding organization had a direct conflict of interest (e.g., when SET corporation funded the study), the funder interest was considered present.

**Department type.** Each author's department was classified as one of the following: corporate unit, evaluation unit, education or psychology department, business or economics

department, and other/unknown.

We examined the effects of publication period, COIs, and authors' department using the random effect model (using restricted maximum-likelihood estimator or REML and Fisher's  $z$  transformation of correlations) with specific moderators. All reported analyses were conducted using R, and more specifically, using packages meta, metafor, and metasens.

## Results

Table 1 shows all 97 multisection studies included in Uttl et al.'s (2017) meta-analysis with each study size ( $n$ ) and SET/learning correlations for instructor ( $r$ ) taken from Uttl et al.'s Table 2. The table includes a column showing the presence or absence of each conflict of interest: corporate (Corp), administrative (Admin), evaluation unit (Eval U.), SET author (SET Auth.), funder (Funder), and the total count of all COIs present for each study (i.e., the sum of corporate, administrative, evaluation unit, SET author, and funder interests). In addition, the table includes a column indicating whether or not authors were from educational psychology or psychology (E/Psy) and from business or economics departments (B/Econ).

Table 2 shows the means,  $SD$ s, and a matrix of the simple correlations among all variables including the COIs, department memberships, SET/learning correlation, study size, and publication period. Notably, these correlations indicate that studies published 1981 and after were associated with lower prevalence of COIs and higher involvement of business and economics departments. Moreover, a corporate COI was positively correlated with all other COIs except the administrative COI. Most critically, the size of SET/learning correlations was associated with publication period (studies published in 1981 and after reported lower SET/learning correlations), and with corporate COIs, SET author COIs, and COI count.

Figure 1, top left panel, shows the magnitude of SET/learning correlations as a function of the multisection study size (Uttl et al., 2017). This figure shows that (1) the number of sections included in multisection studies was generally small with many studies based on as few as five sections, (2) many studies reported impossibly high correlations, and (3) the magnitude of SET/learning correlations decreased for larger sized studies (Uttl et al., 2017). Figure 1, top right panel, shows the histogram of multisection studies by publication year. A large number of studies was published just prior to Cohen's (1981) meta-analysis and a relatively smaller number of studies have been published since then. Figure 1, bottom left panel, shows the number of studies published prior to 1981 and since then by origin of the study. Prior to 1981, most studies were published by authors from SET corporations (Corp), evaluation units (Eval U.), administration (Admin), and education and psychology departments (E/Psy). In contrast, since 1981, most studies were published by authors from business and economics departments (B/Econ) and psychology departments (E/Psy). No studies were published by authors from SET corporations or from evaluation units. Figure 1, bottom right panel, shows the study size as a function of publication year. SET/learning correlations reported by studies published prior to 1981 vs. since then used smaller samples. That is, most studies published prior to 1981 used very small samples and larger samples became more common only after 1981.

Figure 2, top left panel, shows the size of SET/learning correlations as a function of publication date, prior to 1981 vs. 1981 and after. Each dot represents one SET/learning



correlation superimposed on the boxplot of their distribution. Studies published prior to 1981 reported much larger SET/learning correlations than studies published in 1981 and after. The random effect meta-analysis (using z-transformed correlations and REML estimation method) with publication period as a moderator showed significant differences between SET/learning correlations reported by studies published prior to 1981 and studies published 1981 and after,  $Q(1) = 10.31, p = .0013$ . For 69 studies published prior to 1981,  $r = .31$ , 95% CI = (.22, .39),  $I^2 = 53.1\%$  whereas for 28 studies published in 1981 or after,  $r = .06$ , 95% CI = (-.07, .19),  $I^2 = 19.2\%$ .

Figure 2, top right panel, shows boxplots of SET/learning correlations by authors' associations with SET corporations. SET/learning correlations were much larger when at least one author was associated with a SET corporation. The random effect meta-analysis (using z-transformed correlations and REML estimation method) with corporate COI as a moderator showed significant differences between SET/learning correlations reported by studies with vs. without corporate COI,  $Q(1) = 15.38, p < .0001$ . For 15 studies with corporate COI,  $r = .58$ , 95% CI = (.41, .71),  $I^2 = 5.8\%$ , whereas for 82 studies without corporate COI,  $r = .18$ , 95% CI = (.10, .25),  $I^2 = 52.7\%$ .

Figure 2, bottom left panel, shows boxplots of SET/learning correlations by authors' associations with SET corporations, administration, evaluation units, and authoring SET. For these analyses, studies were assigned to categories by the highest, most significant COI present in the following order: corporate, administrative, evaluation unit, SET author, and no identified COI. Corporate COI was associated with the largest effects, administrative COI with the next largest effects, evaluation unit COI with smaller effects, SET author COI with still smaller effects, and the smallest effects were reported by studies with no known COI. The random effect meta-analysis with COI degree as a moderator showed significant differences between the groups of studies,  $Q(1) = 28.54, p < .0001$ . Table 3 highlights that corporate COI resulted in the largest effect of  $r = .57$ , administrative COI in  $r = .33$ , evaluation unit COI in  $r = .25$ , SET author COI in  $r = .15$ , and no known COI in  $r = .06$ . Neither the SET author COI nor the no known COI estimates significantly differed from zero.

Figure 2, bottom right panel, shows boxplots of SET/learning correlations by authors' associations with SET corporations, administration, evaluation units, psychology/education departments, business/economics departments, and other departments. Similarly to the previous analysis, the highest level association was used to categorize studies. The figure highlights that authors associated with education and psychology, and with business and economics, reported similar SET/learning correlations when authors involved in administration and in evaluation units were classified in administration and evaluation units rather than as ordinary members of their departments. The random effect meta-analysis with authors' association as a moderator showed significant differences between the groups of studies,  $Q(1) = 33.82, p < .0001$ . Table 4 highlights that authors associated with education and psychology reported  $r = .15$ , authors associated with business and economics reported  $r = -.04$ , and authors with no known associations reported  $r = .16$ .

Figure 3 shows boxplots of SET/learning correlations by the total number of COIs identified for each study. The figure highlights that studies with no COIs reported on average

nearly zero SET/learning correlations whereas studies with 1, 2, or 3 COIs reported on average small to moderately large SET/learning correlations. The random effect meta-analysis with the number of COIs as a moderator showed significant differences between the groups of studies,  $Q(3) = 22.35, p < .0001$ . Table 5 highlights that studies with no identifiable COI resulted in estimated  $r = .06$  whereas studies with the most COIs resulted in estimated  $r = .53$ .

## Discussion

Researchers with a vested interest in finding large positive SET/learning correlations found, on average, large positive SET/learning correlations. In contrast, researchers with no identifiable COIs found that SET/learning correlations were zero or nearly zero. The largest SET/learning correlations were reported by authors with ties to SET selling corporations. Smaller but still substantial SET/learning correlations were reported by researchers with administrative assignments and with evaluation units/departments responsible for the administration of SETs. Moreover, authors with the most significant COIs were publishing their studies primarily prior to 1981 whereas authors with no or less significant COIs were publishing their studies in 1981 or afterwards. Studies published prior to 1981 reported small but significant ( $r = .31$ ) SET/learning correlations whereas studies published in 1981 and after reported near zero, non-significant SET/learning correlations ( $r = .06$ ). As our analyses show, the presence of COIs was associated with earlier publication date but also with smaller samples. Finally, whereas corporate, administrative, and evaluation unit authors nearly ceased publishing multisection studies on SET/learning correlations, authors from business and economics departments are now responsible for a substantial portion of newer, larger, and higher quality studies published in 1981 and after.

Our findings are striking but at the same time not surprising. As Ahn et al. (2017) recently showed, even effects reported in randomized clinical trials -- a gold standard of experimental design used to determine effectiveness of therapies -- are correlated with authors' conflict of interest. When a principal investigator had financial ties with a drug manufacturer, a study was more than three times as likely to report positive drug effects than when the principal investigator had no such financial ties. Just like drug manufacturers, SET corporations need to sell SETs to as many universities and colleges as they can, as their profits as well as their presidents', scientists', and employees' salaries depend on such sales. SET corporations reporting that SET/learning correlations are low or near zero would terminate their SET related revenues and reasons for their existence. Similarly, evaluation units need to show their value to the administration and reporting that SETs do not correlate with learning may result in the demise of these units. Moreover, as argued previously, administrators need cheap and quick ways to evaluate faculty and if the administrators report that SETs do not correlate with learning they could not justify using them and would have to search for more costly and more demanding alternative ways to evaluate faculty's teaching effectiveness. To start with, they would have to figure out what effective teaching is, something that educators were unable to establish despite years of sustained effort. Finally, the COI influence on SET research findings is likely facilitated by a general lack of transparency and openness. SETs are typically not made public and only SET corporations, evaluation units, and administrators have access to SET data. As a result, the findings originating from these units cannot be examined and verified by independent researchers.

How COI blinds corporate presidents and scientists to facts is clearly demonstrated by the IDEA Center's reaction to Uttl et al.'s (2017) criticism of SETs as a measure of faculty's teaching effectiveness. Shortly after Uttl et al.'s paper was published online (in September 2016) and Inside Higher Education ([insidehighered.org](http://insidehighered.org)) published the story by Colleen Flaherty about its findings, Ryalls (President, IDEA Center), Benton (Senior Research Officer, IDEA Center), and Li (Research Associate, IDEA Center) (2016) criticized Flaherty and Uttl et al. on the IDEA Center website in an opinion piece titled "IDEA Editorial Note #3: Response to 'Zero Correlations Between Evaluations and Learning'". Ryalls, Benton and Li (2016) went as far as to misattribute ideas expressed by Abrami and d'Apollonia (known proponents of SETs as a measure of teaching effectiveness) to Uttl et al. and to misquote Uttl et al. to make their misattribution appear genuine. Table 6 shows what Uttl et al. wrote verbatim, including their quote from Abrami and d'Apollonia (1999), compared to what Ryalls, Benton, and Li stated Uttl et al. wrote. Clearly, Ryalls, Benton, and Li chose to omit our quotation from, and reference to, Abrami and d'Apollonia, misquoted what we wrote, and made it appear as if we held Abrami and d'Apollonia's views that students should determine the academic standards via their SET ratings.

Uttl (personal correspondence, October 19, 2016) pointed out to Ryalls, Benton and Li their misrepresentations, misattributions, and misquotes. Ryalls (personal communication, October 19, 2016) chose to ignore the facts and responded that "We [Ryalls et al.] stand by our entire body of work over 40+ years, including our latest editorial note." Thus, the IDEA Center is willing to counter criticism of SETs as a measure of faculty teaching effectiveness with misrepresentations, misattributions, and misquotes of those who author such criticisms, with at least the silent approval of the IDEA Center board members who were copied on the correspondence.

Setting aside Ryalls, Benton, and Li's (2016) misquotes, misrepresentations, and refusal to correct them, their statements about "this misconception", "the assumption that students are out for the easy 'A'", and the hard working students being insulted by this assumption is a collection of words devoid of any logic -- it sounds good but it makes no sense. First, regarding the misconception, Abrami and d'Apollonia (1999, p. 520) clearly and unambiguously wrote that "[SET] ratings help to identify those instructors who do this well [find the balance between what students might learn and what students are capable of learning]." In other words, the SET ratings determine academic standards. If the ratings are high, the academic standards adopted by a professor are appropriate. In contrast, if the SET ratings are low, academic standards are inappropriate, either too low or too high. We agree with Abrami and d'Apollonia that some students indeed rate professors this way. Second, we do not know what Abrami and d'Apollonia assumed but it is a well-known fact (not an assumption) that *some* "students are out for the easy 'A'" and others are seeking education. Third, hard working students would hardly be insulted by the statement that *some* students are out for an easy A. In this case, just as in logic, it makes all the difference whether the statement about students being out for an easy 'A' refers to none, some, or all students. None and all are definitely false. Some is definitely true. Ryalls, Benton, and Li do not seem to appreciate the difference.

Ryalls, Benton, and Li's (2016) work is peppered with numerous other misleading statements. For example, Ryalls et al. criticized Uttl et al.'s (2017) inclusion of Capozza's (1973)

study that reported a SET/learning correlation of  $r = -.94$  because Uttl et al. “criticized previous meta-analyses for including ‘impossibly high’ correlations”. First, Uttl et al. did not criticize previous meta-analyses for including impossibly high correlations in their sample of multisection studies; Uttl et al. criticized them for “failure to notice impossibly high SET/learning correlations and to adequately consider the negative correlation between the SET/learning correlations and sample size. (p. 24).” Second, Ryalls et al. failed to notice or to acknowledge in their opinion piece that Uttl et al. noticed these impossibly high correlations, conducted outlier analyses, identified Capozza (1973) and Rodin (1972) correlations as outliers, and re-ran their meta-analyses and reported results with and without outlier studies included. Ryalls et al.’s reporting that Uttl et al. included Capozza (1973) in their meta-analyses and failed to exclude it is clearly incorrect.

Given that most authors did not disclose COIs and their COIs had to be determined by searching other publicly available sources, it is likely that we missed some COI ties and, in turn, this may have underestimated the strength of associations between COIs and the reported SET/learning correlations. In future studies, it is desirable that authors declare COIs clearly and with specificity within each study.

Our findings suggests that COIs, including corporate and administrative interests, had a major influence on findings of multisection studies, especially the findings of multisection studies published prior to 1981, prior to Cohen’s (1981) meta-analysis that concluded that SET/learning correlations were substantial. When Cohen (1981) worked on his meta-analysis of SET/learning correlations in multisection studies, he did not notice apparent COIs in many of these studies, and as we noted above, he himself had an interest in demonstrating the validity of SET/learning correlations as a measure of teaching effectiveness. At the time of publication of his meta-analysis, Cohen was already an assistant director of an evaluation unit at Dartmouth College specializing in instructional evaluation research. Cohen’s COI may explain why he did not notice strong evidence of small sample bias (Uttl et al., 2017) and why he disregarded dependence of SET/learning correlations on sample size even though reviewers of his meta-analysis were “concerned that rating/achievement correlations vary according to the number of sections used in the study” (Cohen, 1981, p. 303).

## Conclusions

SET/learning correlations reported by multisection studies vary substantially with COI ties of the authors of these studies. Authors with ties to SET selling corporations, administration, and evaluation units found small to moderately large SET/learning correlations. In contrast, authors with no identifiable COIs found zero or near zero SET/learning correlations. The large variability in findings associated with COIs is troubling. Thousands of faculty members have been evaluated using SETs and many had their careers terminated by SETs. Yet, the evidence is now clear that SETs do not measure teaching effectiveness (Uttl et al., 2017), are influenced by a number of TEIFs not attributable to professors (e.g., student interest, student prior knowledge, subject matter, class size), and are also influenced by a number of faculty attributes -- sex, accent, national origin, beauty/hotness -- that universities are ill-advised to use in high stakes personnel decisions because use of such attributes is at minimum illegal.

Our findings highlight the need for openness and transparency, especially when research is likely to be used to support high stakes personnel decisions. At minimum, the conflict of interest ought to be clearly declared, including ownerships of shares in SET selling corporations and salaries derived from such activities, and data made available for other researchers wishing to verify reported findings. Although one may hope that Cohen's (1981) meta-analysis would not be published today without declaration of COI and without Cohen providing at minimum a list of all studies included in his meta-analysis, the SET/learning correlations he extracted for each study, and the number of sections each SET/learning correlation was based on, this is not guaranteed. As the IDEA Center example above demonstrates, when money talks, even presidents and senior scientists of nonprofit corporations do not seem to mind misrepresenting facts.

Finally, we need to encourage authors with no COI ties to conduct relevant research. As noted above, many newer multisection studies were conducted by authors from business and economics departments and by other authors with no COI ties, and without their contributions we may still erroneously believe that SETs are valid measures of faculty's teaching effectiveness.

452  
453

# References

- Abrami, P. C., Cohen, P. A., & d'Apollonia, S. (1988). Implementation Problems in Meta-Analysis. *Review of Educational Research*, 58(2), 151–179. <https://doi.org/10.2307/1170333>
- Abrami, P. C., & d'Apollonia, S. (1999). Current concerns are past concerns. *American Psychologist*, 54(7), 519–520. <https://doi.org/10.1037/0003-066X.54.7.519>
- Ahn, R., Woodbridge, A., Abraham, A., Saba, S., Korenstein, D., Madden, E., ... Keyhani, S. (2017). Financial ties of principal investigators and randomized controlled trial outcomes: cross sectional study. *BMJ*, 356, i6770. <https://doi.org/10.1136/bmj.i6770>
- Benton, S. L., & Cashin, W. E. (2012). Student ratings of teaching: A summary of research and literature. The IDEA Center.
- Capozza, D. R. (1973). Student evaluations, grades, and learning in economics. *Western Economic Journal*, 11(1), 127.
- Clayson, D. E. (2009). Student Evaluations of Teaching: Are They Related to What Students Learn?: A Meta-Analysis and Review of the Literature. *Journal of Marketing Education*, 31(1), 16–30. <https://doi.org/10.1177/0273475308324086>
- Cohen, P. A. (1981). Student Ratings of Instruction and Student Achievement: A Meta-analysis of Multisection Validity Studies. *Review of Educational Research*, 51(3), 281–309. <https://doi.org/10.3102/00346543051003281>
- Cohen, P. A. (1983). Comment on A Selective Review of the Validity of Student Ratings of Teaching. *The Journal of Higher Education*, 54(4), 448–458. <https://doi.org/10.2307/1981907>
- Davis, B. G. (2009). *Tools for Teaching* (2 edition). San Francisco, CA: Jossey-Bass.
- Feldman, K. A. (1989). The Association between Student Ratings of Specific Instructional Dimensions and Student Achievement: Refining and Extending the Synthesis of Data from Multisection Validity Studies. *Research in Higher Education*, 30(6), 583–645.
- Marsh, H. W. (2007). Students' Evaluations of University Teaching: Dimensionality, Reliability, Validity, Potential Biases and Usefulness. In R. P. Perry & J. C. Smart (Eds.), *The*

*Scholarship of Teaching and Learning in Higher Education: An Evidence-Based*

*Perspective* (pp. 319–383). Dordrecht: Springer Netherlands. [https://doi.org/10.1007/1-4020-5742-3\\_9](https://doi.org/10.1007/1-4020-5742-3_9)

McCallum, L. W. (1984). A meta-analysis of course evaluation data and its use in the tenure decision. *Research in Higher Education*, 21(2), 150–158.  
<https://doi.org/10.1007/BF00975102>

Rodin, M., & Rodin, B. (1972). Student Evaluations of Teachers. *Science*.

Ryalls, K., Benton, S., & Li, D. (2016). IDEA Editorial Note #3: Response to “Zero correlation between evaluations and learning.” The IDEA Center.

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the Validity of Student Evaluation of Teaching: The State of the Art. *Review of Educational Research*, 83(4), 598–642.  
<https://doi.org/10.3102/0034654313496870>

Stark, P., & Freishtat, R. (2014). An Evaluation of Course Evaluations. *ScienceOpen Research*.  
<https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1>

Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty’s teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, 54, 22–42.  
<https://doi.org/10.1016/j.stueduc.2016.08.007>

Wachtel, H. K. (1998). Student Evaluation of College Teaching Effectiveness: a brief review. *Assessment & Evaluation in Higher Education*, 23(2), 191–212.  
<https://doi.org/10.1080/0260293980230207>

Table 1

*Studies included in Uttl et al. 's (2017) meta-analysis with identified COIs and authors' associations with education and psychology departments, and business and economics departments.*

Study	n	r	Corp	Admin	Eval U.	SET Auth.	Funder	COI count	E/Psy	B/con
Beleche.2012	82	0.21	0	0	0	0	0	0	0	1
Benbassat.1981	15	0.18	0	0	0	0	0	0	0	0
Bendig.1953a	5	0.89	0	0	0	0	0	0	1	0
Bendig.1953b	5	-0.80	0	0	0	0	0	0	1	0
Benton.1976	31	0.17	0	0	0	1	0	1	0	0
Bolton.1979	10	0.68	0	0	0	0	0	0	1	0
Braskamp.1979.01	19	0.17	0	1	1	0	0	2	0	0
Braskamp.1979.02	17	0.48	0	1	1	0	0	2	0	0
Bryson.1974	20	0.55	0	0	0	0	0	0	0	0
Capozza.1973	8	-0.94	0	0	0	0	0	0	0	1
Centra.1977.01	7	0.6	1	0	1	1	0	3	0	0
Centra.1977.02	7	0.61	1	0	1	1	0	3	0	0
Centra.1977.03	22	0.64	1	0	1	1	0	3	0	0
Centra.1977.04	13	0.23	1	0	1	1	0	3	0	0
Centra.1977.05	8	0.87	1	0	1	1	0	3	0	0
Centra.1977.06	7	0.58	1	0	1	1	0	3	0	0
Centra.1977.07	8	0.41	1	0	1	1	0	3	0	0
Chase.1979.01	8	0.78	0	0	1	0	0	1	0	0
Chase.1979.02	6	0.19	0	0	1	0	0	1	0	0
Cohen.1970	25	0.42	0	0	0	1	1	2	0	0
Costin.1978.01	25	0.52	0	1	0	0	0	1	1	0
Costin.1978.02	25	0.56	0	1	0	0	0	1	1	0
Costin.1978.03	21	0.46	0	1	0	0	0	1	1	0
Costin.1978.04	25	0.41	0	1	0	0	0	1	1	0
Doyle.1974	12	0.49	0	0	1	1	0	2	0	0
Doyle.1978	10	-0.04	0	0	1	1	0	2	0	0
Drysdale.2010.01	11	0.09	0	0	0	0	0	0	1	0



Drysdale.2010.02	10	-0.02	0	0	0	0	0	0	1	0
Drysdale.2010.03	8	0.64	0	0	0	0	0	0	1	0
Drysdale.2010.04	11	0.03	0	0	0	0	0	0	1	0
Drysdale.2010.05	10	-0.23	0	0	0	0	0	0	1	0
Drysdale.2010.06	12	-0.1	0	0	0	0	0	0	1	0
Drysdale.2010.07	11	0.19	0	0	0	0	0	0	1	0
Drysdale.2010.08	16	0.41	0	0	0	0	0	0	1	0
Drysdale.2010.09	11	0.23	0	0	0	0	0	0	1	0
Elliot.1950	36	0.32	0	0	1	0	0	1	0	0
Ellis.1977	19	0.58	0	0	0	0	0	0	1	0
Endo.1976	5	-0.15	0	0	1	0	0	1	0	0
Fenderson.1997	29	0.09	0	0	0	0	0	0	0	0
Frey.1973.01	8	0.91	1	0	0	1	0	2	0	0
Frey.1973.02	5	0.6	1	0	0	1	0	2	0	0
Frey.1975.01	9	0.81	1	0	0	1	1	3	0	0
Frey.1975.02	12	0.18	1	0	0	1	1	3	0	0
Frey.1975.03	5	0.74	1	0	0	1	1	3	0	0
Frey.1976	7	0.79	1	0	0	1	0	2	1	0
Galbraith.2012a.01	8	0.23	0	0	0	0	0	0	0	1
Galbraith.2012a.02	10	0.32	0	0	0	0	0	0	0	1
Galbraith.2012a.03	12	-0.07	0	0	0	0	0	0	0	1
Galbraith.2012a.04	8	0.31	0	0	0	0	0	0	0	1
Galbraith.2012a.05	8	-0.13	0	0	0	0	0	0	0	1
Galbraith.2012a.06	9	-0.16	0	0	0	0	0	0	0	1
Galbraith.2012a.07	13	0.11	0	0	0	0	0	0	0	1
Galbraith.2012b	5	0.29	0	0	0	0	0	0	0	1
Greenwood.1976	36	-0.11	0	0	0	1	0	1	1	0
Grush.1975	18	0.45	0	1	0	0	0	1	0	0
Hoffman.1978.03	75	0.29	0	0	0	0	0	0	1	0
Koon.1995	36	0.3	0	1	0	1	0	2	1	0
Marsh.1975	18	0.42	1	0	1	1	0	3	0	0

Marsh.1980	31	0.38	1	0	1	1	0	3	0	0
McKeachie.1971.01	34	0.06	0	1	0	1	0	2	1	0
McKeachie.1971.02	32	-0.20	0	1	0	1	0	2	1	0
McKeachie.1971.03	6	0.10	0	1	0	1	0	2	1	0
McKeachie.1971.04	16	0.25	0	1	0	1	0	2	1	0
McKeachie.1971.05	18	0.55	0	1	0	1	0	2	1	0
McKeachie.1978	6	0.20	0	0	1	1	0	2	0	0
Mintzes.1977	25	0.38	0	0	0	0	0	0	0	0
Morgan.1978	5	0.92	0	0	0	0	0	0	0	1
Murdock.1969	6	0.77	0	0	0	0	0	0	1	0
Orpen.1980	10	0.61	0	0	0	0	0	0	0	0
Palmer.1978	14	-0.17	0	0	0	0	0	0	0	1
Prosser.1991	11	-0.42	0	0	1	0	0	1	0	0
Rankin.1965	21	-0.06	0	0	0	0	0	0	1	0
Remmers.1949	53	0.28	0	0	1	1	0	2	0	0
Rodin.1972	12	-0.75	0	0	0	0	0	0	1	0
Sheets.1995.01	58	0.15	0	0	0	0	0	0	0	1
Sheets.1995.02	63	-0.25	0	0	0	0	0	0	0	1
Solomon.1964	24	0.30	0	0	1	0	0	1	0	0
Soper.1973	14	-0.17	0	0	0	0	0	0	0	1
Sullivan.1974.01	14	0.51	0	1	0	0	0	1	0	0
Sullivan.1974.04	9	0.57	0	1	0	0	0	1	0	0
Sullivan.1974.05	9	0.33	0	1	0	0	0	1	0	0
Sullivan.1974.06	16	0.34	0	1	0	0	0	1	0	0
Sullivan.1974.07	8	0.48	0	1	0	0	0	1	0	0
Sullivan.1974.08	6	0.55	0	1	0	0	0	1	0	0
Sullivan.1974.09	8	0.08	0	1	0	0	0	1	0	0
Sullivan.1974.10	14	0.42	0	1	0	0	0	1	0	0
Sullivan.1974.11	6	-0.28	0	1	0	0	0	1	0	0
Sullivan.1974.12	40	0.40	0	1	0	0	0	1	0	0
Turner.1974.01	16	-0.51	0	0	0	0	0	0	0	0

Turner.1974.02	24	-0.41	0	0	0	0	0	0	0	0
Weinberg.2007.01	190	0.04	0	0	0	0	0	0	0	1
Weinberg.2007.02	119	-0.26	0	0	0	0	0	0	0	1
Weinberg.2007.03	85	-0.09	0	0	0	0	0	0	0	1
Whitely.1979.01	5	0.80	0	0	1	1	0	2	0	0
Whitely.1979.02	11	-0.11	0	0	1	1	0	2	0	0
Wiviott.1974	6	-0.04	0	0	0	0	0	0	0	0
Yunker.2003	46	0.19	0	1	0	0	0	1	0	1

---

458

459

460

461 Table 2  
462 Means, standard deviations, and correlations among key variables.  
463

	<i>M</i>	<i>SD</i>	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. <i>n</i>	20.86	25.91										
2. <i>r</i>	.24	.38	-.17									
3. 1981+	.29	.46	<b>.29</b>	<b>-.27</b>								
4. Corp	.15	.36	-.16	<b>.38</b>	<b>-.27</b>							
5. Admin	.25	.43	-.03	.12	<b>-.26</b>	<b>-.25</b>						
6. Eval U.	.24	.43	-.13	.16	<b>-.30</b>	<b>.36</b>	<b>-.21</b>					
7. SET Auth.	.31	.46	-.11	<b>.27</b>	<b>-.38</b>	<b>.64</b>	-.07	<b>.41</b>				
8. Funder	.04	.20	-.07	.16	-.13	<b>.34</b>	-.12	-.12	<b>.31</b>			
9. COI count	.99	1.07	-.18	<b>.39</b>	<b>-.51</b>	<b>.73</b>	.19	<b>.60</b>	<b>.85</b>	<b>.34</b>		
10. E/Psy	.30	.46	-.06	-.03	.08	<b>-.22</b>	.15	<b>-.36</b>	-.05	-.14	<b>-.21</b>	
11. B/Econ	.20	.40	<b>.36</b>	<b>-.28</b>	<b>.55</b>	<b>-.21</b>	<b>-.22</b>	<b>-.28</b>	<b>-.33</b>	-.10	<b>-.44</b>	<b>-.32</b>

464  
465 Note. Correlations with  $p < .05$  are printed in bold.  
466

467 Table 3  
468 *Random effect meta-analysis of SET/learning correlations by authors' associations with the most*  
469 *significant COI: Subgroup results.*

Authors' association	<i>k</i>	<i>r</i>	95% <i>CI</i>	<i>Q</i>	<i>I</i> <sup>2</sup>
Corporate	15	.57	(.41, .70)	14.85	5.8%
Administrative	24	.33	(.21, .45)	21.76	0%
Evaluation unit	12	.25	(.03, .44)	11.53	4.6%
SET author	3	.15	(-.16,.43)	4.18	52.1%
None	43	.06	(-.04, .16)	105.30	60.1%

470

471 Table 4  
472 *Random effect meta-analysis of SET/learning correlations by authors' associations with the most*  
473 *significant unit/department: Subgroup results.*

Authors' association	<i>k</i>	<i>r</i>	95% <i>CI</i>	<i>Q</i>	<i>I</i> <sup>2</sup>
Corporate	15	.57	(.41, .70)	14.85	5.8%
Administrative	24	.33	(.21, .45)	21.76	0%
Evaluation unit	12	.25	(.04, .44)	11.53	4.6%
Education/Psychology	18	.15	(-.02, .31)	36.88	53.9%
Business and economics	18	-.04	(-.18, .10)	39.88	57.4%
Other	10	.16	(-.04, .34)	23.26	61.3%

474  
475

476 Table 5  
477 *Random effect meta-analysis of SET/learning correlations by number of COIs present:*  
478 *Subgroup results.*

Number of COIs	<i>k</i>	<i>r</i>	95% <i>CI</i>	<i>Q</i>	<i>I</i> <sup>2</sup>
0	43	.06	(-.05, .16)	105.30	60.1%
1	24	.33	(.20, .45)	23.88	3.7%
2	18	.30	(.14, .44)	26.25	35.2%
3	12	.53	(.34, .67)	9.50	0%

479  
480  
481

482 Table 6  
 483 A quote from Uttl et al. (2017) compared to Ryall, Benton, and Li's (2016) misrepresentation of it  
 484 in their opinion piece published by the IDEA Center on their website.

**Uttl, White, and Wong Gonzalez (2016, 2017) (the text omitted by Ryalls, Benton and Li in bold)**

**Ryalls, Benton and Li's (2016) writing (p. 5)**

**It has been argued that SETs are responsible for grade inflation and work deflation in higher education by shifting the responsibility for students' learning and grades from students to professors.**

**In response, Abrami and d'Apollonia (1999) opined:**

**" . . . academic standards that are too high may be as detrimental to the learning of students as academic standards that are too low. The art and science of good teaching is finding the balance between what students might learn and what students are capable of learning. We believe that ratings help identify those instructors who do this well." (p. 520)**

**In this view, SETs are some sort of measurement instrument device enabling professors to find what students' perceive to be an appropriate workload and an appropriate amount to learn for specific grades, in short, an appropriate academic standard from students' perspectives.**

**Professors who do this well, argue Abrami and d'Appolonia, will get high SETs. In contrast, professors who are either unable to do it well or do not do it because they believe that such student determined academic standards are detrimental to the students' themselves and/or to the society at large will get poor SETs. It follows that if the student determined standards are too far off from the standard necessary to pass the next course, attain a degree, or succeed in a new career after graduation, a professor is faced with a stark dilemma: teach to the SET and be promoted and tenured, or teach to prepare students for the next course, graduation and future careers, and be terminated.**

A final point we wish to challenge in the Uttl et al. article is their commentary on student-determined academic standards, a commentary which really does not follow from the findings in their study:

"SETs are some sort of measurement instrument device enabling professors to find what students' perceive to be an appropriate workload and an appropriate amount to learn for specific grades, in short, an appropriate academic standard from student's perspectives...professors who are either unable to do it well [i.e., teach to student determined standards] or do not do it because they believe that such student determined academic standards are detrimental to students' themselves and/or to the society at large will get poor SETs (p. 19)."

This misconception has been around since the first time SRI were collected, and, unfortunately, gets repeated frequently along with its corollary that "easier" teachers get higher SRI. The assumption that students are out for the easy "A" is insulting to students who are working hard to gain an education.

485  
 486



# Figure Captions

**Figure 1. Descriptive analyses of SET/learning correlations, sample size, and publication year.** Top left panel shows the magnitude of SET/learning correlations as a function of the multisection study size (see Uttl et al., 2017). Top right panel shows the histogram of multisection studies by year of publication. Bottom left panel shows the number of studies published prior to 1981 and since then by origin of the study. Bottom right panel shows study size as a function of year of publication.

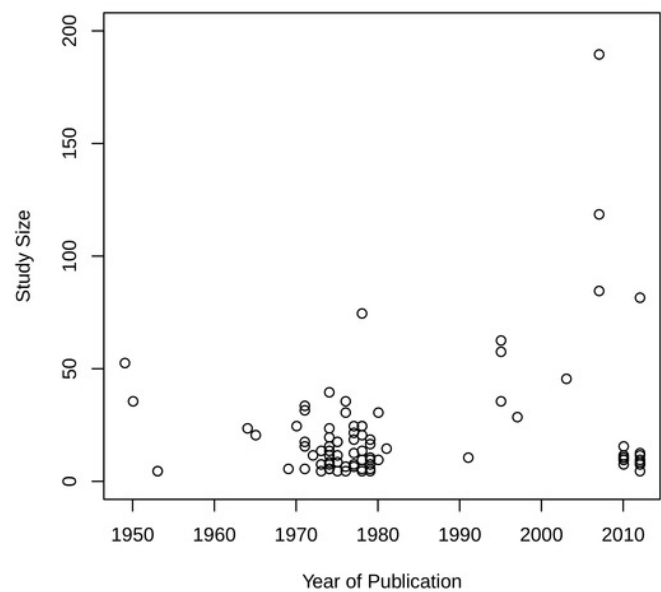
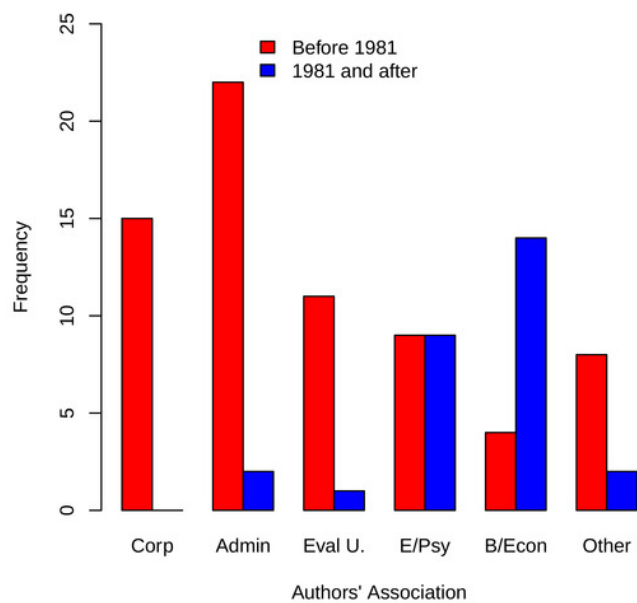
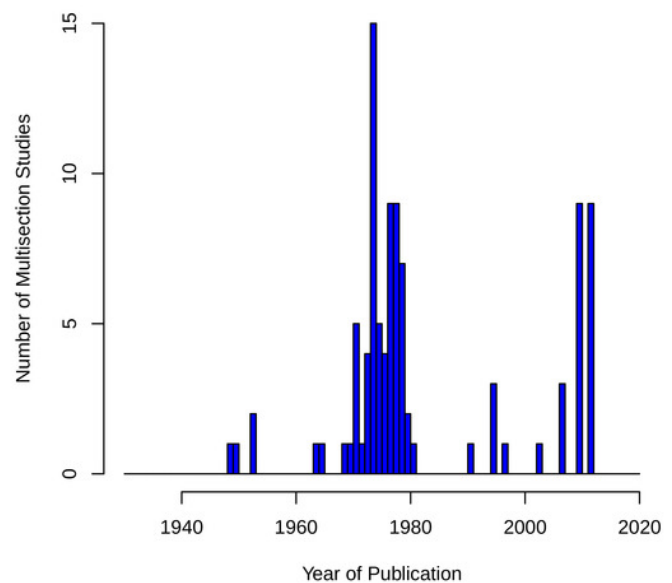
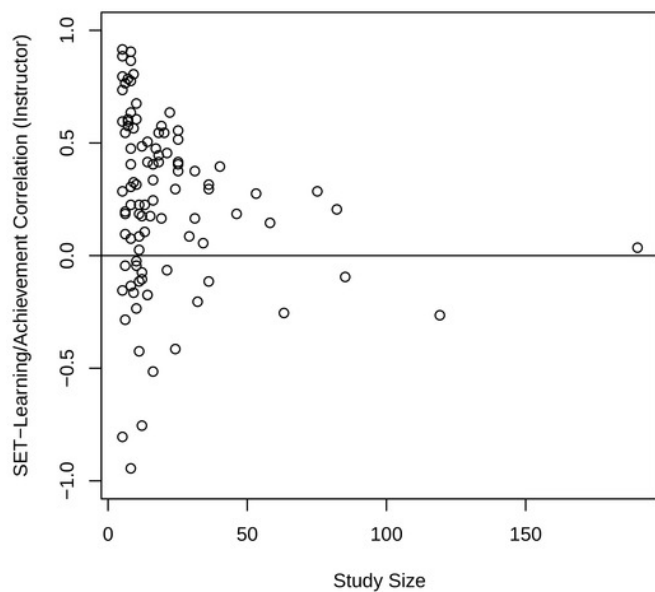
**Figure 2. SET/learning correlations as a function of publication time, COIs, and authors' associations.** Top left panel shows the size of SET/learning correlations as function of publication date, prior to 1981 vs. 1981 and after. Each dot represents one SET/learning correlation superimposed on the boxplot of their distribution. Top right panel shows boxplots of SET/learning correlations by authors' associations with SET corporations. Bottom left panel shows boxplots of SET/learning correlations by authors' associations with SET corporations, administration, evaluation units, and authoring SET. For these analyses, studies were assigned in categories by the highest, most significant COI in the following order: corporate, administrative, evaluation unit, SET author, and no identified COI. Bottom right panel shows boxplots of SET/learning correlations by authors associations with SET corporations, administration, evaluation units, psychology/education departments, business/economics departments, and other departments. Similarly to the previous analysis, the highest level association was used to categorize studies.

**Figure 3. SET/learning correlations by the total number of COIs identified for each study.** The figure shows SET/learning correlations as a function of number of COIs. Each dot represents one SET/learning correlation superimposed on the boxplot of their distribution.

# Figure 1

Descriptive analyses of SET/learning correlations, sample size, and publication year.

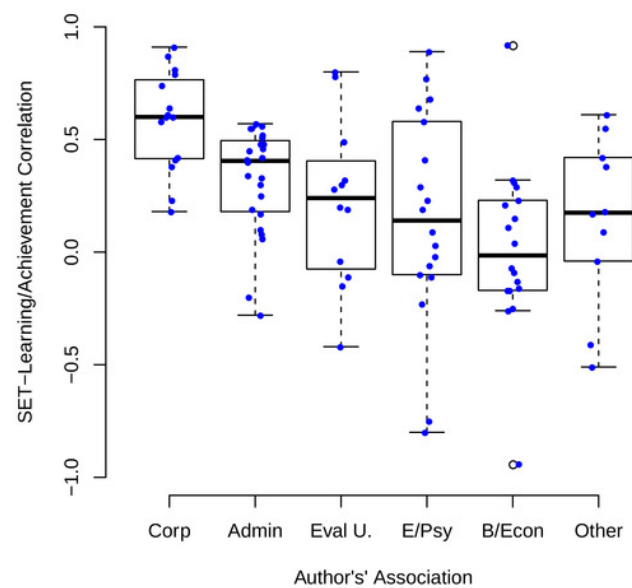
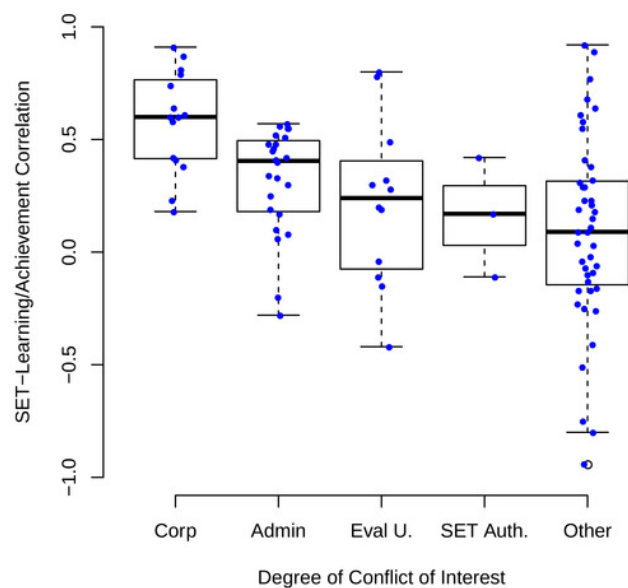
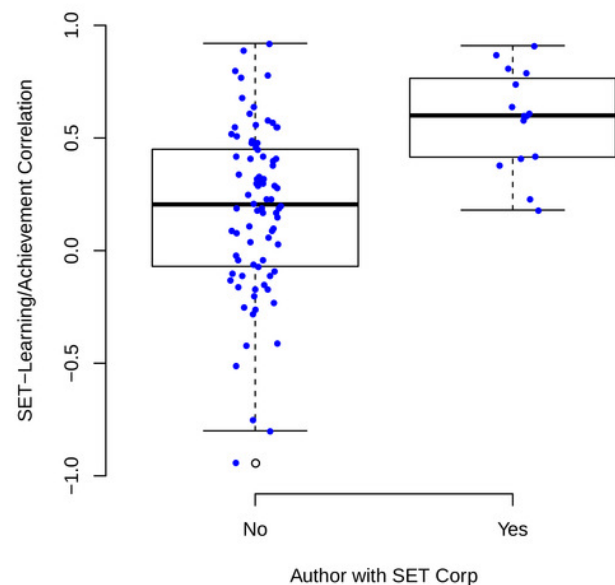
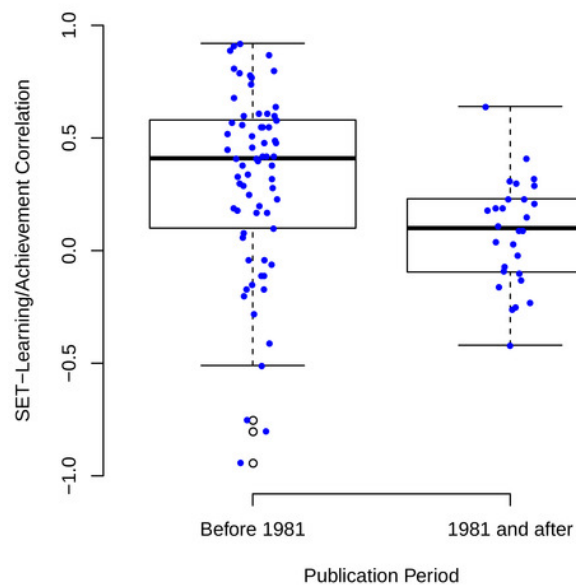
Top left panel shows the magnitude of SET/learning correlations as a function of the multisection study size (see Uttl et al., 2017). Top right panel shows the histogram of multisection studies by year of publication. Bottom left panel shows the number of studies published prior to 1981 and since then by origin of the study. Bottom right panel shows study size as a function of year of publication.



# Figure 2

SE/learning correlations as a function of publication time, COIs, and authors' associations.

Top left panel shows the size of SET/learning correlations as function of publication date, prior to 1981 vs. 1981 and after. Each dot represents one SET/learning correlation superimposed on the boxplot of their distribution. Top right panel shows boxplots of SET/learning correlations by authors' associations with SET corporations. Bottom left panel shows boxplots of SET/learning correlations by authors' associations with SET corporations, administration, evaluation units, and authoring SET. For these analyses, studies were assigned in categories by the highest, most significant COI in the following order: corporate, administrative, evaluation unit, SET author, and no identified COI. Bottom right panel shows boxplots of SET/learning correlations by authors associations with SET corporations, administration, evaluation units, psychology/education departments, business/economics departments, and other departments. Similarly to the previous analysis, the highest level association was used to categorize studies.



# Figure 3

SET/learning correlations by the total number of COIs identified for each study.

The figure shows SET/learning correlations as a function of number of COIs. Each dot represents one SET/learning correlation superimposed on the boxplot of their distribution.

