
education policy analysis archives

A peer-reviewed, independent,
open access, multilingual journal



Arizona State University

Volume 18 Number 14

June 30, 2010

ISSN 1068-2341

Cheating in the first, second, and third degree: Educators' responses to high-stakes testing

Audrey Amrein-Beardsley, Ph.D.

David C. Berliner, Ph.D.

Sharon Rideau, Ed.D.

Arizona State University

Citation: Amrein-Beardsley, A., Berliner, D.C. & Rideau, S. (2010) "Cheating in the first, second, and third degree: Educators' responses to high-stakes testing" *Educational Policy Analysis Archives*, 18 (14). Recovered [date] from <http://epaa.asu.edu/ojs/article/view/714>

Abstract: Educators are under tremendous pressure to ensure that their students perform well on tests. Unfortunately, this pressure has caused some educators to cheat. The purpose of this study was to investigate the types of, and degrees to which, a sample of teachers in Arizona were aware of, or had themselves engaged in test-related cheating practices as a function of the high-stakes testing policies of No Child Left Behind. A near census sample of teachers was surveyed, with valid responses obtained from about 5 percent, totaling just over 3,000 teachers. In addition, one small convenience sample of teachers was interviewed, and another participated in a focus group. Data revealed that cheating occurs and that educators can be quite clever when doing so. But how one defines cheating makes it difficult to quantify the frequency with which educators engage in such practices. Our analysis thus required us to think about a taxonomy of cheating based on the definitions of 1st, 2nd, and 3rd degree offenses in the field of law. These categories were analyzed to help educators better define, and be more aware of others' and their own cheating practices, in an attempt to inform local testing policies and procedures.

Keywords: education policy; high-stakes tests; cheating; validity; unintended effects.

Haciendo trampa en el primer, segundo y tercer grado: Las respuestas de los docentes a las

evaluaciones con consecuencias severas

Resumen: Los educadores están siendo fuertemente presionados para asegurar que sus estudiantes obtengan buenos resultados en las pruebas de rendimiento. Desafortunadamente, esas presiones hacen que algunos educadores hagan trampas. El objetivo de este estudio fue investigar los tipos y grados de conocimiento que educadores en el estado de Arizona tenían sobre o que habían participado directamente en prácticas fraudulentas en relación a las políticas de evaluación de desempeño relacionadas con la ley No Child Left Behind. Una muestra cuasi-censal de profesores con un total de 5% de respuestas válidas y poco más de 3.000 profesores fue investigada. Además, un grupo pequeño (muestra de conveniencia) de los profesores fue entrevistado, mientras que otro grupo participo en un grupo focal. Los datos revelan que existe el fraude y los educadores actúan de manera muy inteligente cuando los resultados. Sin embargo como uno define el fraude dificulta medir cuantitativamente la frecuencia con la cual los profesores tienen este comportamiento. Por esa razón, nuestro análisis requiere pensar en una taxonomía de falsificación basada en las definiciones de 1º, 2º e 3º niveles de infracciones legales. Estas categorías se analizaron para ayudar a los educadores a definir con más precisión y ser más conscientes de los casos de fraude cometidos por otros maestros y por sí mismos, en un intento de reflexionar sobre las políticas y procesos de evaluación.

Palabras clave: política educativa; evaluaciones con consecuencias severas; falsificación; validez; efectos no deseados.

Engano no primeiro, segundo e terceiro grau: as respostas dos professores às avaliações com conseqüências graves

Resumo: Educadores estão sob pressão para garantir que seus alunos obtenham ótimos resultados nos testes de desempenho. Infelizmente, tal pressão abre precedentes para que educadores falsifiquem os resultados. O objetivo deste estudo foi investigar os tipos, além das variações, de uma amostra de professores do estado do Arizona que percebiam ou que estiveram envolvidos em práticas fraudulentas em função das políticas de avaliação de desempenho do programa No Child Left Behind. Uma amostra dos professores foi investigada, com as respostas válidas obtidas a partir de cerca de 5%, totalizando pouco mais de 3.000 professores. Além disso, uma pequena amostra de conveniência de professores foi entrevistada, enquanto outra amostra participou de um grupo focal. Os dados revelaram que as fraudes ocorrem e que os educadores agem de forma inteligente quando falsificam os resultados. Porém, a maneira como cada um define fraude dificulta a possibilidade de mensurar a frequência com que educadores se envolvem em tais práticas. Nossa análise, portanto, exige pensar sobre uma classificação de falsificação baseada nas definições de 1º, 2º e 3º graus de infrações nos domínios da lei. Tais categorias foram analisadas para ajudar os educadores a definir de forma mais precisa e estar mais conscientes das práticas de fraude realizadas por outros professores e por eles mesmos, em uma tentativa de informar as políticas e procedimentos de avaliação local.

Palavras-Chave: políticas educacionais; avaliações com conseqüências graves; falsificação de resultados; validade; efeitos não-intencionais.

Introduction

This study begins in a teachers' lounge during the week of the administration of Arizona's high-stakes test – Arizona's Instrument to Measure Standards (AIMS). A teacher listened as her colleagues discussed tips and tools, explaining how they helped their students do better on that morning's math test. One teacher admitted to giving students rulers so they could more easily convert metric to standard measurements. Another described helping students read the test. And

another admitted to having “always done things like that.” The teacher listening, the teacher-researcher in this study, asked her colleagues if they believed these practices might qualify as acts of “cheating.” They responded with an adamant “No!”

We know that many educators are discouraged by high-stakes testing and fearful of their results – results that are used to make consequential decisions, largely determined by state and federal policies, specifically the stronger accountability policies written into the No Child Left Behind Act of 2001. In effect (Haver, 2004; Nichols, Glass, & Berliner, 2005), not only have educators admitted that they have “cheated” on high-stakes tests, they have acknowledged knowing colleagues who have cheated as well. Share a tale of cheating with another educator and in return expect a “Have I got a story for you!”

Phelps (2005) wrote, “If a teacher’s performance is judged, in whole or in part, on the basis of their students’ test results, certainly they are given an incentive to cheat” (p. 49). Nichols and Berliner (2007) made the case that high-stakes testing almost always corrupts the indicators used and the educators judged by such tests. On the other hand, Sacks (1999) questioned the frequency and prevalence of cheating on high-stakes tests, ultimately arguing that it is “more likely rare than common” (p. 126). Lorrie Shephard (1990) argued nearly a decade earlier that cheating “is generally believed to occur in a very tiny percentage of schools (1–3%)” (p. 20). But both of the latter estimates were calculated pre-NCLB. Just recently analysts found that one in five elementary and middle schools across the state of Georgia submitted “highly abnormal” answer sheets with almost 90% of one school’s scores suspect (Fausset, 2010).

Across our country, newspaper articles continue to be published about teachers and administrators accused of cheating on high-stakes tests (Applebome, 2009; Fausset, 2010; Fessenden, 2007a, b; Gootman, 2008). Yet an accurate number of incidences of cheating is still unknown. We can only assume with confidence that there are many more incidents of teachers and administrators cheating on high-stakes tests than are reported in the media.

We argue that educators are cheating on high-stakes tests more often than most would like to think, especially since the enactment of NCLB (2001). And we know of no scholar who would disagree. The main question that remains, however, is how often cheating occurs. In addition, while sometimes educators are cheating unknowingly (depending on how one defines cheating), we argue that educators are certainly aware of their own or colleagues’ test-related practices which might be considered *questionable* – practices which might be professionally debatable, disputed, and justifiable or indefensible given whether they ultimately support the best interests of students.

However defined, we argue incidences of cheating are bound to be underreported and underestimated. There is a code of silence that exists in our schools where the shared belief is to protect students, teachers, administrators, and schools, ultimately to preserve the very things at stake (e.g., grade-to-grade promotions, high school diplomas, job stability, employment, monetary bonuses and rewards, school accreditation, school takeover or closure, etc.). The professional ethics of our nation’s teachers and administrators are being pushed to a point where their actions have become questionable, embarrassing, and increasingly desperate (Heubert & Hauser, 1999; Johnson, Johnson, Farenga, & Ness, 2007; Jones, Jones, & Hargrove, 2003; Nichols & Berliner, 2007; Orfield & Kornhaber, 2001). And though NCLB may have been written with good intentions, we can declare that its accountability demands and associated sanctions have lead educators to engage more frequently in cheating.

But what is cheating? How one defines cheating is essential when assessing test-related incidences of cheating. And how one defines cheating is fundamental when attempting to make accurate estimations of how common and widespread cheating on high-stakes tests actually is.

Some might define cheating straightforwardly, like when teachers change students' answers on a bubble sheet or teachers provide students with answers either overtly (e.g., reading answer options aloud and orally emphasizing correct choices) or covertly (e.g., talking students through test items with clever hints and suggestions). Similarly, some might accuse administrators of cheating when they change students' answers on a bubble sheet or release secure tests early to give teachers advanced knowledge about the tests.

Others might define cheating more loosely. Perhaps the most common practice in which teachers engage to prepare students for high-stakes tests is teaching to the test. Test officials and educational leaders often dismiss teaching to the test, spinning it as an appropriate test preparation strategy if teachers are teaching to the standards not the test items themselves (see Popham, 2007, for a full discussion of this position). Others might deem either form of teaching to the test as cheating. If teachers are focusing inordinately on the standards they know will be assessed on a high-stakes test, over others, would this not compromise the validity of the test in precisely the same way as teaching to the actual test items? Is this then equivalent to, or a lesser form of cheating, or should such actions be classified as sensible teaching?

If administrators allocate some of their funds to purchase test preparation booklets to help teachers drill students on items like those that are on the test, or promote a narrowing of the curriculum to focus more on the subject areas "that matter" on high-stakes tests, or promote exclusionary practices to stop low-scoring students from taking high-stakes tests, might these be considered forms of cheating?

If administrators hire test consultants to help teachers, and to help teachers help their students game high-stakes tests, is this cheating? The federal government has implicitly promoted this by recently awarding the state of Delaware millions of dollars to employ "Data Coaches." These coaches, part of the Obama administration's Race To The Top (RTTT) program of school funding, can be paid \$100,000 a year. Their sole job is get students to score higher on tests so that school and teacher scores can go up and bonuses for performance might be more easily obtained (Ohanian, 2010). Is this cheating or merely setting conditions for cheating to flourish?

We argue that it is improper and cavalier to think about and define cheating in conventional, clear-cut ways. Indeed all test-preparation practices teachers and administrators use to artificially boost test scores might fall under a better conceptualized taxonomy of cheating. And because practices vary in terms of severity or degree, we turn to the profession of law and its murder statutes to help us further understand the idea of "degrees" of cheating. It must be stated that we are not trying to be sensational or emotive by implying that we think about the felony of murder as analogous to the offense of educators cheating on high-stakes tests. We are simply arguing that how our courts have historically defined degrees of murder might help to inform how we conceptualize the considerably lesser wrongdoing of cheating on high-stakes tests.

Degrees of Cheating

Although murder definitions vary across states, under most statutes murder comes in three degrees. Murder in the first degree is defined as willful and premeditated, often with malice intent, and is often considered the most serious and most worthy of the harshest punishments. The definition of murder in the second degree is more nuanced, pertaining to cases where murder is less intentional and engaged in more casually with more subtlety. Murder in the second degree might not necessarily be intentional, premeditated, or caused by a murderer's lack of concern or compassion for life, but it takes place largely given the circumstances in which it occurs. Judgments and consequences are often based on the likelihood the murder would have occurred in any other similar

situation. And murder in the third degree is usually caused by indifference, recklessness, or negligence, without premeditation or intent, and is often referred to as involuntary manslaughter. There is also the case of justifiable homicide which is defined as a murder that is morally excusable, defensible, or acceptable, enough so that the accused might be exculpated for the murderous act.

If we apply this logic to help us sort out cheating on high-stakes tests, we might be better able to classify and qualify the severity of particular test-related acts. This might help teachers and educators understand that there are lesser degrees of cheating in which they might be engaged without knowing that they are doing so. This might also help educational leaders implement local testing policies to better manage and prevent the incidences of unprofessional or unethical test-related practices.

Cheating in the first degree might be defined as willful and premeditated, and it might often be considered the most serious and most worthy of the harshest punishments. Cheating in the second degree, again, requires finer distinctions because it is often engaged in with more subtlety. As such, it might be defined more casually as it might not necessarily be intentional, premeditated, or caused by an educator's lack of concern or compassion for students, but it might take place largely given the circumstances in which it occurs. Judgments about these types of cheating might be based on the likelihood the act of cheating would have occurred in any other similar situation. And cheating in the third degree might be caused by indifference, recklessness, or negligence, without premeditation or intent, and might also be referred to as involuntary cheating. There may also be cases of justifiable cheating. We will address this category later.

The purpose of this study was to investigate the types of, and degrees to which a volunteer sample of teachers in Arizona were aware of, or had themselves engaged in cheating. As the study progressed, however, we acknowledged that understanding the meaning of cheating was more complex than what many assume, and we recognized the need for an enhanced definition of what it means to cheat. That said, we used our aforementioned taxonomy of cheating to frame this study and the results presented herein.

Literature Review

Instruction based disproportionately on what students need to learn in order to do well on high-stakes tests pollutes the validity of the inferences one might make about what students know and can do (See, for example, Haladyna, 2002; Nichols & Berliner, 2007; Popham, 2007). When faced with pressure to perform at a high level, in any situation, a person may behave inappropriately and engage in practices that ordinarily are not typical of that person. In the case of high-stakes testing, when measuring an outcome (e.g., school quality) with an indicator (e.g., test scores) and instituting a negative consequence (e.g., state takeover) or reward (e.g., a salary increase) as a result of the indicator, this can only lead to the corruption (Nichols & Berliner, 2007) of the measurement system (e.g., high-stakes testing) and people (e.g., teachers and administrators) who work within the system. This corruption comes in varying degrees.

Cheating in the First Degree

Cheating in the first degree might be defined as willful and premeditated, and it might often be considered the most serious and most worthy of the harshest punishments. If educators cheat in the first degree to reap the positive or evade the negative consequences attached to high-stakes tests, they might be deserving of the strongest of reprimands.

Teachers across the country have been charged, formally and informally, with erasing and changing students' test answers to raise their classroom's composite test scores. Administrators have been charged with the same, erasing and changing answers once students' test responses are in

administrative hands. Teachers have been charged with giving students correct answers on tests during class (e.g., pointing to correct answer options), as have administrators outside of the classroom, for example, calling students to the principal's office to help them change responses. Administrators have also been accused of changing student identification numbers so low-scoring students' test-scores are eliminated from results, due to being invalid.

Both teachers and administrators have been accused of not testing academically weak students and engaging in creative practices to exclude students with histories of poor academic performance from participating in tests. Walter Haney (2000) termed the resultant increases in test scores *illusions* arising from student *exclusions*. Educators realized that by excluding students from high-stakes tests, especially students with disabilities and non-English speakers, scores would increase, albeit artificially. In true Machiavellian form (Radin, 2000), the ends justify the means if one defines the ends as the positive consequences reaped, or the negative consequences evaded. Related, educators have been charged with convincing low-scoring students to stay home on test days or high-scoring students to be sure not to miss school, suspending students on testing days, and expelling students before high-stakes tests are administered to make certain such low scoring students would not contribute their low test scores to classroom, school, or district composite statistics.

In general, students who do not do well on high-stakes tests, or are not *test ready* (Haladyna, Nolen, & Haas, 1991), are not wanted when these tests are actually administered. Thus, when they do not want the scores of low-performing students to negatively skew classroom and school composite scores, administrators and teachers have engaged in the aforementioned practices, or simply taken things into their own hands as discussed (Bass & Dizon, 2006; Booher-Jennings, 2006; Fessenden, 2007a, b; Gootman, 2008; Haladyna et al, 1991; Haney, 2000; Nichols, Glass & Berliner, 2006; Nichols & Berliner, 2007; Orfield & Kornhaber, 2001; Sociology of Education Research Group, 1998; Urdan & Paris, 1994).

Cheating in the Second Degree

The pressures associated with high-stakes tests also lead educators to cheat in more subtle ways. Cheating in the second degree might be defined more casually as it might not necessarily be intentional, premeditated, or caused by an educator's lack of concern or compassion for students, but it might occur largely given the circumstances in which cheating occurs. Judgments about these types of cheating might be based on the likelihood the act of cheating would have occurred in any other similar situation.

Educators, often in attempts to help their students, engage in practices casually, knowing that what they are doing might be considered cheating. But they employ such practices nonetheless. They might do so to support their students sometimes given the high-stakes situations in which they are situated.

For example, educators have been accused of encouraging students to redo problems as students are taking tests. An educator who is walking around the classroom and sees that a student made a *stupid mistake*, is surely hard pressed to not hint to the student to check the answer again, many times knowing the student knows better or knew better when they made the incorrect mark. Educators have also been charged with distributing tip (or cheat) sheets that include vocabulary words, equation keys, and the like to provide students with additional resources that might help them answer questions correctly. Then there are the teachers who have been accused of leaving classroom resources around (e.g., rulers, dictionaries) or have kept learning resources visible (e.g., multiplication and alphabet charts) within the region of the room.

Educators might coach students during tests, for example, by providing students with hints about certain test items or talking students through test items and processes, providing them with the tips and suggestions they have been trained to offer under all other circumstances. Or they might be accused of reading questions to students when not allowed to do so, defining difficult vocabulary words, helping students make sense out of questions, or translating test directions, item prompts, or response options for students who might still be learning the language in which the test is written. There are also reported incidences of educators giving students extra time to complete tests, after time is up or during lunch, recess, or special subjects.

While the aforementioned cases of cheating in the second degree are less obvious, largely as they occur within a particular set of circumstances, they still might be considered cheating to a lesser degree. Such actions might not necessarily be intentional, premeditated, or caused by an educator's lack of concern or compassion for students. Indeed the reverse might often be true (Bass & Dizon, 2006; Booher-Jennings, 2006; Haney, 2000; Kelleher, 1999; Nichols et al, 2006; Nichols & Berliner, 2007; McGill–Franzen & Allington, 1993; May, 2000; Olson, 2000; Orfield & Kornhaber, 2001; Schrag, 2000; Sociology of Education Research Group, 1998; Urdan & Paris, 1994).

Cheating in the Third Degree

High-stakes tests also compel educators to cheat to even lesser degrees. Cheating in the third degree might be caused by indifference, recklessness, or negligence, without premeditation or intent. Such incidences might also be referred to as involuntary cheating.

Teaching to the test is arguably the most popular practice in which educators engage to raise scores on high-stakes tests. When educators teach to the test, the standards and/or the actual test items being measured on the tests determine what is taught to students. And as educators become more familiar with the tests, they use what they learn to give their students more of an edge on the tests, that is, until the next version of the test is developed and administered. Teaching to the test comes in many forms, however.

Educators who might have administered past iterations of a high-stakes test could use the information garnered to better prepare students for upcoming, similar tests. They might focus inordinately on what they anticipate in hopes that many of the items and concepts tested will be the same, and all the while believing that this is good test preparation practice. Educators might take a set of testing blueprints, readily available nowadays online via state department of education's websites, and focus only on the *testable* standards, concept areas, and performance objectives, as conveniently classified. Educators might use these as tools to help them focus their instruction, all the while excluding or postponing the other things students should know and be able to do.

Educators might develop series of clone items, or items that look just like the items expected on upcoming tests, with the names of the people changed in the word problems and the numbers in the mathematical items altered, and they might use these to test students repeatedly in preparation for upcoming tests. Or they may have students write and re-write persuasive essays, excluding other writing genres because the other genres matter less, or are less valued on high-stakes tests. They might also make copies of actual tests and test items, or tests and items used previously, to rehearse students for upcoming tests, over-and-over again. Or they might have students spend hours memorizing facts, using multiple-choice answers to solve mathematical problems backwards, eliminating unlikely item distractors, making educated guesses, learning test taking strategies, bubbling score sheets, and the like (Berliner, 2009; Cimbricz, 2002; Firestone, Camilli, Yurecko, Monfils, & Mayrowetz, 2000; Heubert & Hauser, 1999; Linn, 2000; Linn, Graue & Sanders, 1990; McNeil, 2000; Madaus, West, Harmon, Lomax, & Viator, 1992; Madaus & Clarke, 2001; Sacks, 1999; Smith & Fey, 2000; Thompson, 2001; Urdan & Paris, 1994).

Cheating in the third degree might also capture other types of cheating. Narrowing the curriculum, for example, occurs when educators neglect the subjects about which they are indifferent, or are less concerned, during high-stakes tests. That is, because certain subject areas do not *count* on high-stakes tests, educators oft neglect to teach non-tested subject areas for a certain period time, usually before high-stakes tests are administered. For example, a school administrator may cut art, music, or physical education for a period of time to allow for more focus on math, language arts, and science. Or an elementary school teacher may replace science with mathematics and social studies with language arts to intensify math, reading, and writing instruction, for example. An educator might also eliminate recess to provide amplified opportunities for students to rehearse the basic subject areas tested (Dorn, 1998; Koretz, 1996; Kreitzer, Madaus & Haney, 1989; McNeil, 2000; Sacks, 1999; Swope & Miner, 2000)

In effect, students might learn to correctly answer test questions without entirely understanding the concepts behind their answers. Frequently, however, educators engage in such practices not fully understanding that they may be compromising their students' opportunities to learn more important things than those that matter on high-stakes tests. Under the guise of good test preparation, they isolate the content and concepts they know will be measured and marginalize the other parts of the curriculum that students should know and be able to do.

That said, cheating in the third degree might be caused by indifference, recklessness, or negligence, without premeditation or intent. Such incidences might also be referred to as involuntary cheating. And while some school personnel continue to perpetuate these practices without thinking or paying due attention to what they might be jeopardizing in terms of their students' learning, and often considering such practices as smart teaching, others might consider these practices as cheating, or cheating to the slightest degree of the three.

Methods

Again, the purpose of this study was to investigate the types of, and degrees to which a volunteer sample of teachers in Arizona were aware of, or had themselves engaged in cheating. Researchers addressed two questions: (1) In what questionable testing or test preparation practices have educators' colleagues engaged? and (2) In what questionable testing or test preparation practices have educators engaged themselves? Researchers did not question whether the practices occurred because of or with greater frequency since the state's post-NCLB, high-stakes testing policies were implemented, however, and can only argue that incidences such as those described have likely increased since NCLB was passed.

Data Collection

Using a pragmatic, mixed methods approach, researchers collected data to better comprehend and expand understandings about these test-related behaviors, collecting data from conducting an online survey, individual interviews, and a focus group (Creswell, 2008; Day, Sammons, & Gu, 2008; Ercikan & Wolff-Michael, 2009; Flowers, 2010; Greene, Caracelli, & Graham, 1989; Johnson & Onwuegbuzie, 2004; Miretzky, 2009).

Survey. The survey instrument (see Appendix A) consisted of three sections with 12 total items. The survey was piloted with 12 educators in the state who belonged to a professional literacy organization. Once revised, it was distributed statewide, via Zoomerang.com (Market Tools, Inc., 2008). The final survey instrument consisted of three sections. The first included eight demographic items (e.g., school types, Title 1 classifications, NCLB performance levels, and respondent demographics), the second included nineteen binary items (yes, no) meant to capture respondents' knowledge about the test-related behaviors of their colleagues, and the third included the same

nineteen binary items, included in the final section to capture respondents' self-reported behaviors. Two free-response items were also included to allow for elaboration.

To measure the instrument's levels of internal reliability, researchers collapsed the nineteen binary items into three constructs, as aligned with the taxonomy and defined in degrees of first, second, and third degree cheating. They categorized the first seven items in the first degree, the next seven items in the second degree, and the last five items in the third degree (see Appendix A). Researchers then measured internal consistency reliability using Cronbach's alpha. Each of the combined sections had a moderately high alpha coefficient (first = 0.70; second = 0.67; third = 0.83) indicating that the survey was reliable. The analysis resulted in alpha values close to or higher than the generally accepted 0.70 level (Cronbach, 1951). But it should also be noted that the instrument included only dichotomous, binary variables, which yielded lower bound, or underestimates of the instrument's reliability (Sun, Chou, Stacy, Ma, Unger, & Gallaher, 2007).

Interviews. The teacher-researcher involved in this study also conducted unstructured, individual interviews one-on-one wherever nine participants who volunteered to be interviewed felt comfortable discussing this topic. Interviewees were invited on the basis of the personal relationships they had with the teacher-researcher – the aforementioned teacher who overheard her colleagues discussing their test-related practices in the teachers' lounge. Interviewees completed a consent form granting researchers the permission to record and use their information as part of the data collected. Participants had the option of not participating or withdrawing at any time. The teacher-researcher started the interviews by asking participants to share their stories about cheating on high-stakes tests. The natural discussion was used to determine the interview trajectory from there. Interviewees were likely more apt to volunteer if they had stories to tell, impacting subsequent levels of generalizability. As such, the forthcoming interview data should be interpreted accordingly.

Focus group. Data were also collected from a focus group. Participants were invited by letter, again, on the basis of their personal relationship with the teacher-researcher conducting the focus groups, again, warranting caution in interpretation. Upon arrival, participants completed a consent form giving their permission to record and use the shared information for dissemination purposes. They also signed a confidentiality agreement. Again, participants had the option of not participating or withdrawing at any time. The teacher-researcher asked participants to share their experiences about educators cheating. The focus group participants lead the discussions from there.

Sample

Survey. Arizona educators (including teachers and administrators) were solicited to respond to an online survey on which they were asked to anonymously and voluntarily report their own and others' questionable test-related practices; the latter might be considered hearsay. Researchers purchased the email addresses for 59,597 educators in the state from an online marketing company. Researchers sent the survey to all email addresses included on the list, once initially and again six days after the survey was launched. The rationale for emailing such a large number of educators was the belief that the number of responses might be low, which it was.

Twelve percent (7,412) of the emails were returned due to invalid addresses. The instrument remained open for 15 days, and when researchers closed the survey 3,085 educators had responded. This yielded a five percent response rate including all email addresses purchased and a six percent response rate including only the email addresses considered valid.

With a volunteer sample such as this, which is not random or representative, sample size is usually not relevant. Although the sample of respondents is large, as a percentage of all of the educators emailed it is small, and as a volunteer sample no statistical generalizations are possible. It is also likely that participants who had strong feelings about high-stakes tests responded at greater

rates, and participants who did not respond did not, biasing the sample further. The results from this sample of respondents, therefore, are of dubious generalizability and should be interpreted with caution. Nevertheless, they still deliver a strong message about cheating under high-stakes testing policies.

The majority of the respondents (71%) were teachers, who represented various grade levels, only 4% of whom reported no teaching experience. They were not removed from the sample, however, in that they still had stories of interest to tell. They had experience being class monitors and observers, serving in peripheral roles within schools. Conversely, the plurality (43%) of respondents reported 10 or fewer years of teaching experience. The majority (60%) of respondents had master's degrees and 5% had earned their doctorates. The others (35%) had not earned a degree beyond a bachelor's.

Ninety-five percent of the respondents reported that they worked in the public school system, with an additional 3% coming from charter schools. Forty-eight percent of the schools were located in suburban areas, 26% in urban areas, and 26% in rural areas. Approximately half (47%) were Title 1 schools. Using Arizona's NCLB classification system, thirty-one percent of the schools were labeled as excelling, 18% were highly performing, 17% were performing plus, 18% were performing, 7% were under performing, 1% was failing, and 8% of the respondents were unaware of their school's label.

The data obtained, however, do not reflect or capture where, and the extent to which cheating occurs. This is because many respondents stated that they had engaged in such practices or heard about cheating elsewhere or in a prior position. And again, researchers did not question whether the practices occurred because of or with greater frequency since the state's post-NCLB, high-stakes testing policies were implemented.

Interviews. In terms of the interviews, two interviewees were spouses and were interviewed together, yielding eight total interviews. The interviews included (1) a Latina teacher who formerly taught in a high-needs school, (2) a female principal who happened to be at the same school, (3) a Latina teacher in a low-needs school district, (4) a Caucasian, married couple both of whom taught at a high-needs school and one of whom was honored as an exemplary teacher by a local foundation, (5) a Caucasian, male teacher at a high-needs school, (6) a Latina teacher at a magnet school who had been formally accused of cheating, (7) an African-American, female, reading coach at a high-needs school, and (8) a Caucasian, female teacher within the same high-needs district.

Focus group. In terms of the focus group, eight female teachers aged 32 to 60, with 7 to 15 years of teaching experience, participated in the focus group. Again, participants were invited by letter, on the basis of their personal relationship with the teacher-researcher conducting the focus groups.

Data Analyses

With the survey data, researchers analyzed the binary responses converting into percentages the numbers of respondents who acknowledged that they were aware of others' cheating practices or cheated themselves. Percentages illustrating these data are illustrated in a frequency chart included in the findings (see Figure 1). Means, standard deviations, and effect sizes were not calculated as they are generally not very useful for nominal variables (Barrett, Gloeckner, Leech, & Morgan, 2006; Diamantopoulus & Schlegelmilch, 2006). Participants' responses to the open-ended, free-response questions included on the survey instrument were read, coded, and then collapsed and organized into overall themes (Miles & Huberman, 1994).

In terms of the data collected during the interview and focus groups, with intent to supplement the survey results with more in-depth, case stories, researchers analyzed content by

systematically identifying the common and special characteristics of participants' stories (Berg, 2004) in a cross-case analytic form whereby each case was compared and analyzed with respect to other cases (Miles & Huberman, 1994).

Results

Cheating – Survey Data

The first section of the survey instrument included questions about respondents' awareness of others' questionable test-related practices. Respondents were questioned about whether they were aware of colleagues who had engaged in a list of practices that might be considered cheating, before, during, or after the administration of high-stakes tests. In the next section researchers invited participants to respond to the same questions but specifically requested information regarding whether they themselves had participated in the same cheating practices (see Figure 1). Note that when researchers designed and piloted the survey instrument, researchers were unaware of all of the cheating practices in which educators may have engaged; hence, they requested that participants tell their stories in addition to responding to the dichotomous list of questions.

In descending order, respondents reported knowing of situations in which their colleagues encouraged students to redo test problems (39%), gave their classes extra time to complete tests (34%), wrote down questions to help prepare students for future tests (24%), wrote down vocabulary words for the same purposes (23%), and read questions aloud when not allowed to do so (23%). They reported least often knowing others who changed student identification numbers on test booklets to eliminate low scores (1%) and encouraged others to cheat (3%).

Respondents who participated in this study knew of few colleagues (10%) who outright cheated, or cheated in the first degree. Respondents knew more of their colleagues who engaged in second (21%) and third degree (18%) cheating practices, in that order. In other words, teachers were substantially more likely to engage in subtle and casual than negligent cheating practices versus what we have classified as the most unprofessional and unethical test-related practices – cheating in the first degree.

Respondents reported engaging in the following practices themselves, most to least: encouraging students to redo problems (16%), giving students extra time on the tests (14%), coaching students during tests (8%), and leaving materials on the walls during testing (8%). They reported least often changing student identification numbers in order to eliminate low-scoring students' results (0%).

Very few of the educators who participated in the survey admitted to outright cheating, or cheating in the first degree (1%). Respondents were substantially more likely to admit that they also engaged in second degree (7%) than third degree (4%) cheating practices, again in that order. In other words, teachers were substantially more likely to admit to engaging in casual than negligent cheating practices versus what we have classified as the most unprofessional and unethical test-related practices. From these data, it seems that the two most common questionable test-related practices in which educators may engage involve encouraging students to redo test problems and giving students extra time to complete tests.

The biggest threat to the reliability of these results might come from social desirability bias. That respondents might not respond in truthful ways at the risk of being unfavorably viewed by others seemed to have caused inconsistencies in two ways. First, respondents knew of others who

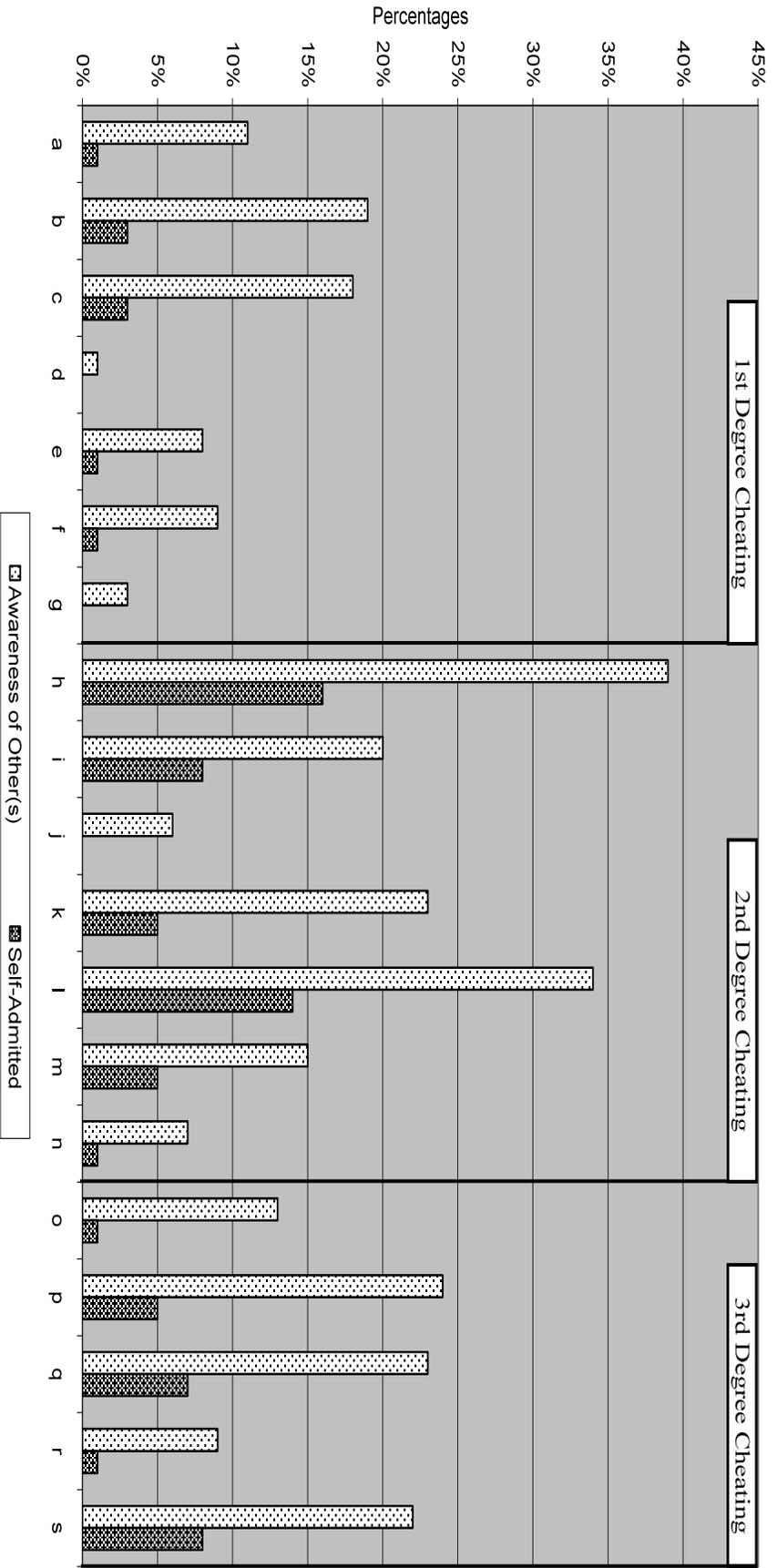


Figure 1. Respondents' awareness of other(s)' and self-admitted cheating practices
 Cheating Behavior Key: a. Erased and changed test answers; b. Gave answers to students; c. Pointed to correct answers; d. Changed student identification numbers in order to eliminate low-scoring students' results; e. Not tested academically weak students; f. Asked students to be absent during testing days; g. Encouraged others to cheat; h. Encouraged students to redo problems; i. Coached students during the test; j. Created cheat sheets for students to help prepare them for tests; k. Read questions to students when not allowed to do so; l. Gave students extra time on tests; m. Allowed students to work on tests during lunch, recess, or specials areas; n. Allowed students to figure out answers on scratch paper for them to check prior to having students mark answer sheets; o. Made copies of tests in order to teach to the tests; p. Wrote down questions to prepare for following years' tests; q. Wrote down vocabulary words to teach to following years' classes; r. Kept test booklets in order to familiarize students with the tests; s. Left materials on the walls knowing it would help students.

had engaged in questionable test preparation practices but did not either admit to, or substantially under-reported engaging in the same practices themselves. For example, respondents noted most often that they were aware of others who encouraged students to redo problems. Yet respondents admitted to the same practices by less than half. Even with anonymity and confidentiality secured, when respondents are solicited to admit to engaging in questionable practices, it seems less likely they will respond truthfully fearing repercussions.

Another threat to validity might be categorized as hearsay. Hearsay occurs when one person is stating a *fact* about another, although the person did not witness the act, nor can they attest to it themselves. In courts, hearsay is commonly dismissed on the basis that it is not a valid indicator of truth. Hearsay may only be admissible when one has a discussion about an event, or practice in this case. Unless educators discuss with one another their questionable, test-related practices, which is not likely, especially if they know what they are doing is wrong, this limits how much respondents really know about what happens in their colleagues' classrooms. As illustrated in Figure 1, it was much more common that respondents heard about these things occurring in others' classrooms, but they claimed not to have engaged in the practices themselves. That said, and as illustrated, the reliability of these data is uncertain.

While these inconsistencies are most likely due to the volunteer, non-random sample, however, the practices about which they were aware and in which they engaged were similar ($r = 0.93$). We believe that this can serve as a loose indicator of the most and least frequent questionable test-related practices in which educators engage. Although respondents may have underreported their own practices, overestimated others, or likely did both, there still existed some underlying structure to their responses. We know of no other data like these.

Respondents were also asked to elaborate on these or other similar incidences. Almost half (43%) made at least one comment. While in this section, a few people responded that cheating does not happen and questioned the intent of the researchers to investigate such nonsense, the plurality of those who provided additional information confirmed that indeed questionable practices occur. The written comments the researchers considered most confirming follow:

- I wandered by a testing room while on break. The teacher said, "So you can see why D can't be the correct answer; it has to be A."
- Another teacher on the team kept bragging about how smart her class was, how her students always scored the best, and how she would have the highest scores again...At test time she encouraged me to "clean-up the test."
- I have had an administrator tell low level students not to come to school.
- When I taught in an elementary school in this area, the administrators sent the test booklets back to us and had us erase the incorrect answer, teach the kids the problems, and help them "fix" the mistake.
- I received several directives to change answers. I refused and the request was repeated several times. I refused each time and finally I was not asked again.
- A first grade teacher told me years ago that she had taken the SAT9 tests home and looked through her students' answer sheets. She told me, "I knew that they knew the answers but they got them wrong on the test. I changed all the answers."
- A teacher was writing answers on the board and I told the principal and nothing ever happened because they did not want the district to look bad. It makes it really hard to stay honest and true when even your boss turns the other cheek to rule breakers!

- I knew a teacher who would tap her finger on the answer or make a face while walking around if an answer was wrong. She also put things into her own words a lot, instead of making sure to follow what the teacher book said.
- The ONLY time I ever encountered ANY type of “cheating” by teachers on a standardized test, was when I was teaching in an extremely poor, inner city school. The school was facing sanctions, and a FEW teachers helped their students in an effort to secure proper funding. It’s sad that the schools which need funding most are the ones we take money away from. While the actions of those teachers were unethical, they only did it to secure funding for their school and students.

In terms of expanding on their own practices, less than half (17%) of those who made comments about others’ practices (43%) wrote at least one comment. Again, the issue of social desirability likely impacted what respondents were willing to share about themselves. And again, respondents questioned the intent of the researchers or responded that while it may occur elsewhere, they do not nor have they ever cheated. The written comments in support of this case follow:

- I have thought about [cheating] every time I give a test, but I have to live with myself.
- I maintain a strict code of ethics in all I do.
- I would never do such a thing.
- If you truly do your job, there is no reason to cheat.
- I am enraged when others tell me about how they have “unconsciously” assisted their student[s].
- Everybody gets screwed. The teacher that cheats looks good, and the honest ones that come after have to pick up the pieces and try to educate students that were cheated.
- I wished I had.
- It drastically challenges the integrity of what we are here for. We are here to guide students to grow and become successful within their individual abilities. Cheating for them cheats them from realizing and utilizing their abilities to become successful. Not all of them are going to be businessmen, attorneys and legislators, so cheating won’t be a successful skill for them to learn.

Otherwise, respondents confirmed that they too had cheated in one sense or another.

- I have checked students work and given answers.
- I have read questions to students vocally emphasizing key terms.
- I admit that I have encouraged students to redo a problem. I have also reworded problems that I felt were worded poorly.
- I gave the students the definition of the words.
- I keep tests from year to year to get a general idea of what is to come.
- Knowing what is going to be on the test to teach to it is the whole idea.
- I remembered the vocabulary words after administering the test. I made a mental note of words like spool and lemonade that I knew my students were not familiar with and incorporated these words in my teachings for the next year. I still teach spool and lemonade even though those words are not on the test anymore.

- The student did not speak the language so it was very unfair to take a test that one could not read. So I said stay home in the morning and come after lunch. I allowed the student to do the math section because he could do the problems.
- Testing is a stressful time. I do not give answers. I simply remind the students of things that would allow them to retrace the connection in their mind to help them find an answer they already have.
- It's hard when you see them going too fast or making silly mistakes to just sit there and not help them in some way. I want them to learn... I don't want these students to feel like they are less of a person because some stupid test that tests one week of their year.
- I felt it was unfair to ask students who knew how to solve a problem, but didn't understand the question, to get an answer wrong. The test didn't test what they knew, just what they could understand of the directions or the wording of the problem.

Respondents took issue with one item included in this section of the questionnaire. They did not believe that leaving posters up on classroom walls, or not covering them, or leaving other resources around should be considered cheating at all. Comments representing this belief follow:

- I have not always agreed that we should take down materials that have been there the entire year as “good teaching practices” to be used as references when working in class.
- Covering up materials such as number lines and calendars is time consuming and idiotic.
- Frankly, if a child had the smarts to look up and read it...than kudos to their observation and reading skills. I never tell them to look around the room for clues.
- If a child is bright enough to use a number line, good for that child!
- I am an English teacher and I have materials on my wall for instruction purposes. I have a list of transition words and the writing process on my walls. I don't think this is cheating or giving students an unfair advantage. It simply reminds them about what they already know but might forget in a stressful situation.

Cheating – Interview Data

Interviews were also conducted to capture accounts of test-related cheating. Interviewees were selected if they had stories to tell regarding their experiences with others' or their own cheating on high-stakes tests. The very fact that they agreed to be interviewed meant that they had a story to tell. Readers should interpret their stories as such, as purely anecdotal.

Interview 1

The following came from an interview with a young specialist in gifted education working in a suburban school district. She shared an experience regarding erasing and changing answers on tests that happened to her when she was a student. The school she attended was in a large, urban district with a student population consisting primarily of Hispanic children. The participant volunteered to share the following:

I remember taking [a test] and my teacher asked me to stay in at lunch time to help her. She said she needed help and it was not unusual because I did it all the time. So, I went in and she was sitting there at one of her student desks. She had a pile of booklets next to her on one side and then right in front of her she had the [answer] forms, you know that we had to use. So, it was no big deal. I walked in and she was looking at each one of them erasing things. I just noticed she was doing that. She

told me to sit down. She said that she needed my help because some of the bubbles were wrong and they needed to be fixed. I did not even think to question her. I was just, like, "Okay." So, I sat down next to her and I saw her. She was going down looking at all the bubbles. She told me that some of them were wrong. So, she gave me a stack of them and she just told me this one is supposed to be B or this one is supposed to be C and whatever. So, I just sat there and we fixed them because they were wrong. But I didn't think they were wrong like as in wrong answers. I just thought, "Oh, maybe they're –" I don't know what I thought. I don't even think I questioned why, she just told me to help her and I helped her.

The researcher who conducted the interviews asked whether the interviewee thought this teacher was a good teacher, and she replied:

At the time I think I thought she was a really good teacher. Even with my teaching experience, I thought back to what she did in the classroom and I thought she was a really good teacher...I think that's why I would do whatever she wanted me to do.

The interviewer asked her what she thought were the effects of this teacher's behavior, and she replied:

I think it affected me personally. I can remember walking into the classroom very vividly. I think that even though at the time I didn't question it, I think instinctively when I was sitting there erasing, I knew it was wrong and that is why I can remember it so well. It was just a retarded test, no big deal, but I remember the sheet, the form, maybe I questioned it in my head, but I did not question her. I think it affected me. I knew that it had. I put this teacher on a pedestal because she was so different, because she was teaching us so many different things. Yet, I was cheating. Why would I be erasing this?

While discussing this teacher the interviewee became teary-eyed and admitted to cheating on the tests as a teacher herself. She stated that she did not think of it as cheating at the time. She had a few English Language Learners in her classroom that began to cry because they could not complete the test. She stated that it broke her heart to see them upset, so she read the test to them.

Interview 2

The next set of stories came from an interview with a retired principal. She had been a teacher for many years and then moved into administration. At the time of the incidents, the interviewee was a principal at a Title 1 school in an urban district.

One bilingual 4th-grade teacher had six of her monolingual Spanish speakers score 96%-98% on their first [test]. When I got the scores back I was shocked by the unusually high number of students scoring in the 90th percentile. I notified the district office of my concern, and the school district attorney, and I decided that getting copies of the score sheets from the department of education would help in determining what happened.

I called the teacher into my office. Before I passed out the results I asked her how she felt her students did on the test. (I had taped the conversation with her consent as she stated she had nothing to hide.) She replied that she thought that even though it was the first time for some of the students that she thought they did about average. I showed her the scores and asked her if she could explain the high number of designated English Second Language students with such high scores. She said, “Well, I wanted them to have a positive experience.”

I asked her if she altered the test results in any way and she stated bluntly, “No.” I asked her if she deviated from the prescribed dialogue given in the directions and again she replied she hadn’t. She then asked me if I was accusing her of cheating. I replied that I just wanted to know why her students had scored so outstandingly. She replied that it was because she was such a good teacher.

Two weeks later when we received copies of the score sheets, the resulting analysis from the state department verified that answers had been changed and remarked correctly with a different slant. The school district attorney met with me, the teacher, and her representative. After I stated the reason for the meeting the teacher told her representative that she would not be needed. After the representative left the room, the district attorney showed her the state department copies and the findings. She was sufficiently embarrassed and repeated, “I just wanted them to have a positive experience.” She was given 10 days suspension without pay, and she accepted it without a word to anyone.

The second incident occurred as follows:

Another time I walked into the classroom of a teacher that [sic] had not turned in the tests on time to find her erasing the answers. I immediately took all of the test booklets, went to district office, and told an assistant superintendent what I had found happening. She issued me new 2nd-grade booklets and I tested the children myself. The teacher was also given five days without pay, and she did not contest the consequences of her actions.

The last incident occurred as follows:

One day I walked into the teachers’ lounge and found an instructional aide [photo] copying the test. I asked who she was copying it for and she said the 3rd-grade. I asked her if she had copied any other pages similar to these and she said that this was the first batch. I took the booklet and the copies from her and went to speak with the teachers at that grade level. I informed them that what they were doing was illegal and that any further cheating on their part would be considered insubordination and that consequences would be time off without pay. One teacher started to cry, saying, “I told you it wasn’t right and that she would get mad.” The grade-level chair stated indignantly, “How are we supposed to make sure they do well on the test if we don’t practice them on the test?” As a result of the confrontation, I had to monitor all copies made by these teachers.

She added:

In all of the above cases the teachers stated that they just wanted the children to do better on the test... A couple of teachers later admitted that they feared for their jobs, but only after they had gotten caught. I did speak individually with the grade level teachers and one of the five said she went along because she was afraid that if her scores were significantly lower than the rest of the grade level she thought she would lose her job. She was a first-year teacher; her teammates, however, were seasoned veterans who should have known better.

Interview 3

The following interview took place at a restaurant. The researcher interviewed a veteran teacher. The researcher had met the teacher at a department of education workshop, explained the focus of this study, and the teacher volunteered to share her stories, presented as follows:

It wasn't until the mid-nineties when a lot of the accountability became so visible and teachers were under pressure, and that pressure was nothing compared to what it is now. It was probably in the mid-nineties when I was [at] a Kindergarten through 6th grade elementary school located in east Phoenix [60% Caucasian to 40% Latino/a]. I was teaching 5th grade at the time.

We got together as a grade level and testing was in about two weeks and we were brainstorming what we were going to do to get the kids ready. At the time we had something called *Scoring High* and so we would drill the kids constantly. Drill, drill, drill!

Testing came and everybody started testing their kids. We got together as a team and [a] teacher and I started to chat. She was really stressed and she said, "I don't think my kids are going to do well and I just don't know what to do." That is when she told me that the day before during lunch time that she had changed the answers. I did not know what to do with that information because she was telling me in confidence. I was thinking, like, "Don't tell me, why are you telling me?" I think she just wanted to get it off of her chest.

The interviewee then shared another similar experience:

I moved to another school within the same district. It had a more relaxed climate. We knew that data was [sic] important, the test scores were important, but our principal made us feel that we were important, the kids were important, and what's more important were the people at the school not the numbers we bring [sic] in. So, my team and I, we were still in 5th grade, felt that somebody's going to look after us. Somebody genuinely cares about what we feel, you know, the pressure we were going through due to the accountability factor which was really hitting hard.

We would still get together for our team meetings and it wasn't until we got together off campus that the people started to open up. You know, "I have a friend that works at such-and-such school in the same district where the principal was not like

ours.” I remember at this discussion we were having one of the teachers said that she had gotten some M&Ms – the kind sold during the holidays that are green and red. She said that before she came to the school where we were now, she had used the M&Ms as the kids were taking their test. If she put a red one on their desk it meant to stop and do it again, go back, and check it out...I mean, it was a creative idea to say the least. There were a lot of honest teachers that were saying they had done it [cheated] but they had to do it because they had to make sure that some of the scores were going to be higher than normal.

She added:

It is all about the numbers, and we lose sight of the art side of teaching. I think what is happening in our society right now with NCLB and the high-stakes test scores is that we are so focused on the science part of teaching, we have lost the balance. There is also an art side of teaching and that is about the relationships that we forge with our students. We are constantly drilling them and telling them that we have to score high. I think they [cheating educators] feel backed up against the wall. Maybe teachers do not buy into the fact that “On this particular day, this is who you are, kid.”

Interview 4

The next interview was conducted with a married couple of teachers, both of whom volunteered to share their experiences:

They [test administrators] used to hand out the test booklets approximately two to two-and-a-half weeks before the test date. The teachers would all have the test booklets and they were able to look at them and that way they were able to prepare the students for the test with the language, vocabulary, everything... We would sit for hours at a time taking and examining the test questions, probably mimicking exactly what was on the test. I did change everything so that it wasn't truly cheating, but at least the kids knew what the format was going to be. I would not give them exactly what was on the test... You give them the actual format but you change the answers.

One added:

Things happen like this, I am sure, across the country. I am sure it happens everywhere. I think there is so much pressure put on the teachers. Your job hinges on your test score. We had pressure in 1993 but not like it is now. It is different now. I am scared... I remember one year I had a bright class [about whom] I had a feeling that I was going to be questioned regarding their test scores because I knew they were going to do well. So, I had proctors in my classroom during the whole testing time. The next year, I was questioned, but I had covered myself by having proctors in the classroom.

Interview 5

This next educator the researcher interviewed was a 3rd-grade teacher in an inner city school, with a minority-majority population of Latino/a students living below the poverty level. The experiences this teacher shared were as follows:

I observed one time during the AIMS a fellow teacher in the same grade level doing the writing, which is supposed to be done by the children – the prewriting, first draft, final draft. I walked by her classroom and the door was opened. She had the kids lined up at her desk... She told me they were just doing the AIMS test and she was correcting their writing and sending them back to their seat[s] to correct their writing. Of course that year she got kudos because her kids reached the 85th percentile. So, she got great results. She was doing the writing for them, but she got all the pats on the back and everything else. I even brought it to the principal's attention. He did not really want to hear about it because of course if our scores look better they [administrators] don't care. They turn their heads.

Another time I walked into the classroom, and a teacher in 4th grade had an equation on the board. She even told me it was a duplicate of what was on the test but with different numbers. There were other things she had written on the board and she stated that she left it on the board while they were taking the test. No one monitored that, so it was no big deal.

I had a test question one year and it used a pictograph. It had all this information, a whole page. Finally, after reading all of this information about the graph, at the very bottom of the page in small letters it read, "Each picture equals two." It was almost hidden. If the kid got tired of reading they would not even read that part. Almost every one of my kids was missing it, and I have to be honest, I told my kids to stop and read the whole thing. I did not tell them what to look for – just to make sure that they read the entire page. Then they went back and did it and all of them got it right. They do that purposely to trick the kids. The test should be simplistic and straight forward to see if the kids have the skills.

We had a teacher who did workshops for test preparation. I used the test-taking skill of picking the same letter when you don't know the answer, and even that is misleading. Is that cheating? We are told to do it. Is it really giving us a true indication of what their knowledge is or what they have learned?

He added:

These are great teachers – I am not saying they are not good teachers, but I am saying they are getting results on the test that make me look like I am not as good a teacher. I believe that the teachers in my current school do not cheat. Most of them are really good teachers. Otherwise we would not be as low as we are.

The researcher asked the participant, "If they are such great teachers, what do you think motivates them to cheat?"

Just the fact that we are held accountable and no one wants to look bad. I have a class this year that is very low. Six kids at grade level out of 26. I can just about stake my life that my test scores are going to be really bad this year, but what am I going to do about it? I mean, maybe they will do better than I think. So, last year I had really good scores, and this year I am going to have bad scores. All of a sudden, in one year, I become a bad teacher?

They [the students] come to me ill-prepared, and what am I going to do? I can only do so much, but I am still going to be held accountable. I have had kids come to my class, and they are at the 70th percentile in 2nd grade. Then, in my class, 3rd grade, they are at the 20th-30th percentile. So, it makes me look like I really dropped the ball... I know the teachers have cheated when they [the students] come to me at the 70th percentile, and yet they can't read.

I am a single parent, so, hell yeah, I have the motivation to want to have good scores. I don't want to lose that money. I have a kid to raise. I think that we need to be held accountable, but when you tie our money to that [test results] and every set of kids you get is different, it is so frustrating... I was told by a principal I get all the problem children, and I get all the low kids because they know that I will work with them. Is that fair? [And] the principals don't care. If you can do it and get away with it, they want you to do it. I don't care what anyone says. I have gone and talked to two principals in the past about it, and they don't want to hear it.

Interview 6

The next interview took place in the classroom of a junior high teacher located in an urban school district. She had been in the teaching profession for about 29 years and had no knowledge of others' cheating, until the past seven to eight years, "as soon as we went into teach to the test mode."

I have had an experience where I actually caught someone cheating, not on the benchmark, but on the "Big Kahuna" AIMS test, which is the one we have to secure. We have to make sure we follow protocol and there are certain things that teachers can do, and we are instructed to do. One of these things is to make sure the answer sheets are clean and that they are filled in correctly. I walked into her class to ask her a question before testing started, but she had already begun the procedure and she was telling the class that if they did not know the answer to a question to leave it blank. I did not think anything of it... "Only answer the ones that you know." So, when I came in, she was bubbling in the empties with the answer sheet of the smartest kid in the class, and she just copied those answers to the empty bubbles on the test. She was pretending to be cleaning up the test. I asked her, "What are you doing?" and she said she was just filling in the circles because they were not filled in properly. She asked me not to mention it to anyone. I didn't.

I have seen other teachers walking around the classroom as the kids are testing point to the answer. I have never pointed to an answer, but I have seen that they had a bad answer, and I have suggested that the student might want to check and to remember what we learned about this and that. I guess this is a form of cheating, too.

The researcher asked her what she believed motivated teachers to cheat. Her response was as follows:

Money. We as teachers live at the poverty level. We are way down there in income. They legislate monies for us and what do they do? They make us jump through hoops for it. They say they are going to raise salaries and make our job comparable with other professions, but bologna, it hasn't happened. The district gets the money, they make themselves top heavy, and we are left in the dark jumping through hoops just to get the money that was legislated to us.

Interview 7

This interview was conducted in the office of a reading coach who had been employed by the same district, in a Title 1 school, for 25 years. This is where the incidents she describes occurred.

We had two teachers at this one particular school that [sic] worked together. One was a Teach for America teacher and the other was a seasoned teacher. The seasoned teacher had been in trouble before by another principal for cheating and was suspended for two weeks... Somehow, she managed to get the younger teacher to pull the test off of a website. They took that test and they actually taught the test to the students. Due to the fact that we were a Reading First school, a person from the department of education would come through to see if we were following the implementation checklist... She pulled a couple of the kids aside and started testing them too. As these two students were walking, they asked, "Is this the one about the ice cream?"

I remember a 1st-grade teacher who used to complain all the time, "My kids can't read. They can't read." My instructional aide was the proctor assigned to her room. He told me after the test that she had read that test verbatim, word for word. Those kids scored almost 100%, and she never got called on the carpet... She still works for the district.

Interview 8

The final interviewee had been teaching for over 35 years in the same district at the same school in central Phoenix. She observed a teacher who "put all of the answers on the board (A, B, C, etc.)" after which "all of the students aced the test." She also explained the relationship between her school and their feeder school (Kindergarten through 3rd grade). The school across the street managed to be labeled performing plus, when her school was labeled under-performing. She explained:

They probably are not [a performing plus school]. They [educators] probably cheated. I see the children and they come here... Somebody did some fancy footwork with those tests because there is no way that these children were E'ing [Excelling] and M'ing [Meeting Standards]. I bet you when we take the test this year they are going to say, "Oh, look how they went down when they went over to the other school." I know that there is no way that these children scored that high. We see them. We see them.

She continued:

I think the ones that [sic] cheat are probably the ones that [sic] are not very good teachers. They have probably not done what they were supposed to do in the classroom. They really haven't taught. They hand out worksheets and there is no teaching going on. They know that they can't teach.

In sum, the cheating behaviors observed and experienced, as detailed by the interviewees, were similar to the test-related behaviors identified in the survey. Across both data sources researchers found that teachers engaged in similar practices (e.g., changing or filling in blank bubbles on answer sheets, correcting student work, pointing to correct answers, reading tests aloud, prodding students to double-check incorrect items, making copies of tests, distributing tests early to allow for advance preparation, using clone items to more exactly prepare students). And they did so for similar reasons (e.g., fear of losing their jobs, associated bonuses and rewards, pressure by colleagues or administrators, feeling of competition, pressure to maintain high-scoring reputation, professional insecurities about student performance, wanting students to succeed).

Cheating – Focus Group Data

One focus group was also formed, and eight educators who all taught in the same district were invited to participate. The purpose of the focus group was to also allow participants to share their stories in a safe setting with social facilitation. After their stories were told these participants volunteered their thoughts about what may motivate educators to cheat.

In terms of personal accounts, participants described more stories, each one still unique but clearly familiar:

- I would say I never intentionally cheated on a test while giving one, but I am not going to say I haven't looked at a test previously. I looked through it to see what was on it...I have never changed answers or done anything like that, but I know people who have.
- We were given a directive in my last school district that if a child asked a question or asked for an explanation, you could do it out loud for everyone to hear and that was okay. So if someone asked, I would stop the test, and I would read that question and answer the student's question so that everyone had the same information.
- [Our] lead 2nd-grade teacher [knew] there was a question [on the upcoming test] and the answer was blizzard, she made sure to tell us, and encouraged us to teach the word blizzard to our student[s] because we are in Arizona so our students do not know the definition of this word. Is this interpreted as cheating?
- One of my best friends taught right next to me and she would go through the book and ma[k]e sure that her kids knew the information before passing out the test... Did I tell my friend, "You should not be looking at the test?" No, I did not. Do we have a code of silence?
- There was a principal in another district and she went to the teachers and handed them a test and told them to use this with their kids, but if anyone asked they did not get it from her.
- Another stated that she did not know of anyone who cheated on the test, nor has she ever done so. She stated that it had never even crossed her mind to cheat: Her "kids either got it or they didn't."

In terms of motivating factors, focus group participants offered up the following:

- “What is the motivation for principals to make sure this doesn’t happen?” Not much, because they do not have a continuing contract they are judged primarily on their school label.
- Teachers cheat because... it is not what you do with those kids, it is not how much those kids value being in your classroom. It is, “what did your kids score at the end of the year?”
- [The motivator] is public humiliation. At my last school district this principal would actually copy and post the teacher’s name and their scores after it. The top teacher in every grade level would get rewarded and the rest of us would feel like complete idiots... This information [was] also shared with [the] PTSO [Parent Teacher School Organization] and site council. So, then the parents want their kids in certain classes and teachers are pitted against one another.
- I think a lot of people fear the repercussions. What will happen to me?
- It comes down to Maslow’s Hierarchy of Needs: Survival. If it comes down to money, a lot of people [would] do the same thing.
- Because we know that even the worst teacher puts up with extreme situations in the classroom, and we know that we don’t get paid enough.
- I think the public has been brainwashed to only look at the score. And that is the way you judge education, and that is the way you judge a teacher.

In sum, seven of the eight educators in the focus group shared more than one story about educators they knew who had cheated on high-stakes tests. The stories told were also similar to those collected via the survey and interviews. A new behavior a focus group member shared dealt with a teacher clarifying questions aloud with entire classes of students, and this allegedly occurred as per district administrator recommendations.

Motivations for cheating were also similar to what researchers found in the data collected via the survey and interviews. In general, the focus group believed that humiliation, reward, competitiveness, reputation, and survival were all key factors motivating teachers to cheat. The group also agreed that cheating is becoming more acceptable and teachers might be increasingly more willing to cheat, and also to believe that it is okay to do so. Of interest, the group also expressed that there did not seem to be much motivation for school administrators to stop cheating from occurring. They believed that when cheating was stumbled upon or reported it was often ignored, but if not, the consequences for cheating were negligible.

Findings

Our asking educators to talk about cheating and share their experiences contributes to the extant literature. But the quantification of the rate of these practices continues to elude us. In this volunteer, non-random sample, more than 50% of the respondents reported having known colleagues who had cheated, and more than 50% reported having engaged in these practices. These figures are certainly not trivial.

Although we did not question whether the practices occurred because of or with greater frequency post-NCLB, and we can only argue that incidences such as those described have increased since NCLB was passed, what we evidence is that educators are cheating at non-trivial rates especially if one defines cheating more loosely, including what we define as cheating in the second and third degrees, in nuanced/subtle/casual and in negligent ways.

The rates of cheating we present are based on data gathered from a volunteer sample of teachers in our survey, but the practices of which they were aware and in which they engaged were

similar ($r = 0.93$). Because there existed some underlying structure to their responses, these data may serve as a reasonable indicator (and the best of which we are currently aware) of the most and least frequent questionable test-related practices educators employ to help their students pass high-stakes tests (see Figure 1).

From participant interviews and focus group discussions, we present data evidencing that when pressured to do well on high-stakes tests, educators engage in quite clever practices, largely to protect themselves and their students. Each of the participants involved had more than one incident to share regarding their colleagues' or their own cheating practices. Participants also expressed their frustrations with high-stakes tests and the state's policies designed to hold them accountable for what they supposedly were not doing as educators.

What seems to be considerably less common, however, are educators cheating in the first degree, in intentional and premeditated ways. This is heartening and a tribute to the teachers and administrators of the nation who have recently had more pressure exerted on them to score ever higher on tests than at any other time in their professional lives. Nevertheless, the negative effects on validity are essentially the same regardless of whether cheating in the first, second, or third degree takes place.

Strong consequences give rise to questionable test-related practices, putting into question the validity of the inferences we make based on such tests (Haladyna et al, 1991; Nichols & Berliner, 2007). If the numbers of incidents of cheating were small, validity would probably be sufficient for the purposes for which the tests were designed. But if cheating occurs at higher rates, than the inferences made when interpreting test results are problematic. This is the conundrum that our educational policymakers must face. A non-trivial contaminant has entered into the process whereby inferences are made about whether schools are "good" and "bad." And the accuracy of the inferences made about the value-added by teachers to students and classes over time, using these same test scores, is being further reduced to yet another unknown degree (Amrein-Beardsley, 2008).

One result of this study might be to help educators better define and be more aware of others' and their own cheating practices, in an attempt to inform local testing policies and procedures, and ultimately to restore the levels validity currently being compromised. We defined cheating in the first degree as willful and premeditated, and considered it the most serious and most worthy of the harshest punishments. We defined cheating in the second degree more casually as it might not necessarily be intentional, premeditated, or caused by an educator's lack of concern or compassion for students. And we defined cheating in the third degree as caused by indifference, recklessness, or negligence, without premeditation or intent.

We also addressed the issue of whether the practices engaged in by many teachers engage might be considered cheating at all. While one educator might consider some of these practices smart, others might consider them unprofessional or unethical. And some might not even consider them at all, unless they are prodded to reflect on what it is they are doing in preparation for, and on high-stakes tests, and they are asked to reflect on whether their test-related, instructional practices ultimately support the best interests of their students.

Professional development promoting discussions of these issues might remind teachers and administrators of their professional obligations. If teachers and administrators are educated about how to prepare and test their students appropriately, in order to secure the validity of the inferences made about educational quality using high-stakes test scores, such ill effects of high-stakes testing policies might better be negated, and abated. Local testing policies, perhaps designed by local educational leaders, might then help control or curb the incidences of artificially raising test scores via cheating. But of course that will only be the case if the stakes are not so high as to make teachers and administrators either ignore moral and professional norms, or choose to fight inappropriate

policy with equally inappropriate behavior. The latter case raises the issue of justifiable cheating, the fourth of the classifications we introduced earlier.

Justifiable cheating is defended by some educators. Teachers and administrators apparently have employed some of the cheating practices noted above to reconcile their beliefs about the proper role of educators in the face of the injustice perceived to be associated with contemporary high-stakes testing policies. Justifiable homicide may be excused or forgiven, given the circumstances in which a murder takes place. In war, for example, soldiers are excused from murdering those with whom they are in combat. People who kill when they are defending themselves or their property might be vindicated after doing so. An offensive act may also be understandable if it was needed in order to prevent or protect others from greater harm.

This brings up the question of whether the questionable practices in which educators engage, especially in the second and third degrees, could ever be justifiable. Should teachers be excused for some of their questionable practices because they are in combat, against what they might properly consider their enemy? Should teachers who cheat be excused if they believe they are protecting their bilingual, special education, or culturally diverse students? Should teachers who cheat be spared because they consider their acts as a necessary response to politicians and business leaders who seem not to understand the goals of a public education system?

Nichols and Berliner (2007) noted that some forms of resistance to high-stakes testing by professional educators include cheating. They report on educators who see high-stakes testing as immoral and unprofessional and undermine the validity as an act of professional necessity. In addition, being put into such difficult moral situations increases the incidences in which teachers must question their own professional integrity (Mehrens, 1998; Sheldon & Biddle, 1998). It is not surprising that many of them feel they must combat injustice with injustice.

The following comments from our respondents best capture these sentiments:

- When our pay is involved, how can we not try to prepare the students as much as possible?
- I don't think teachers would do this if the nation weren't so obsessed with test scores and the media feeding frenzy that comes with reporting the scores, getting the labels, sending letters home to parents admitting shame, embarrassment, failure. If the sanctions were removed, the behaviors would go away.
- Can you blame the little people for trying to work and live in a broken system when those in charge of fixing the system don't really give a frick?
- Usually I would say that any form of cheating is wrong. I consider myself an ethical person. I would say that blatantly giving students answers, changing scores, or telling kids to stay home is definitely wrong. But I have a hard time allowing student learning to be measured by badly made tests that don't even reflect what they are supposed to be evaluating, so that's when I don't feel that I'm cheating by giving kids help on an answer.
- I think that most teachers don't start out thinking, "I'm going to cheat on the test so my students will do well and I'll look good." I think when you walk past a child and look down and see them choose a wrong answer to a question that you KNOW that student should be able to answer correctly, it sometimes "overtakes" you.

This makes it quite difficult to give a simple answer to the question: When high-stakes tests lead professionals to believe that what happens in schools is wrong, to what degree is cheating justified? Are there instances where, like justifiable homicide, teachers and administrators might be excused from cheating given the unreasonable or threatening conditions in which they are situated?

Conclusions

Sacks (1999) stated that, “Just how prevalent cheating is on achievement testing in American schools is probably impossible to quantify, and it’s more likely rare than common” (p.126). As Sachs noted, it would indeed be difficult to quantify the frequency of cheating due to the sensitive nature of the topic. Present is the fear of admitting to something that is professionally and ethically wrong, no matter how one might justify their actions. Yet now also present are more external pressures to cheat as a function of the stronger accountability system associated with NCLB and the high-stakes testing policies associated with that law.

Surely there is nothing new about cheating, however. Such behavior has likely been around from the dawn of mankind and even exists among other species in the animal kingdom. Callahan (2004, p. 14), the author of *The Cheating Culture*, states “Cheating is everywhere.” But as common as it is now and was in the past, cheating is not condoned by the vast majority of people. Paradoxically, although we are a country that frowns upon the lack of values and morality when cheating occurs, we have also become a country in which cheating occurs in all walks of life. Moreover, it is not only often ignored, it is often justified as acceptable practice. Sadly, Callahan (2004) states that “available evidence strongly suggests that Americans are not only cheating more in many areas but are also feeling less guilty about it” (p. 13).

But educators’ cheating on high-stakes tests still strikes most people as unimaginable. Teachers, a bit like clergy, are held to high moral standards and are expected to act as role models for children. Perhaps that is the reason why educators’ cheating on high-stakes tests strikes most people as unbelievable. In the public’s eye educators should be among the last to engage in cheating. But the conditions under which educators now work, and the high-stakes testing policies by which they are now controlled, have pressured more educators into doing so. That some defend their actions as morally justified suggests that we need ethicists as much as we need measurement specialists for the interpretation of the data obtained.

Our conclusion is this: Policies that clearly undermine the moral and professional behavior of America’s teachers need to be debated more thoroughly, and such policies must be challenged if their negative effects outweigh their positive effects on the educational system of our nation. It serves no one’s interest to have policies that inherently promote cheating, and even justifications for cheating by educators, because the policy environment in which they work has become so onerous. There are better ways to design accountability systems.

References

- Amrein-Beardsley, A. (2008, March). Methodological concerns about the Education Value-Added Assessment System (EVAAS). *Educational Researcher*, 37(2), 65-75.
- Applebome, P. (2009, March 7). Grades fixed: An allegation shocks no one. *The New York Times*. Retrieved August 15, 2009 from http://www.nytimes.com/2009/03/08/nyregion/08towns.html?_r=1
- Barrett, K. C., Gloeckner, G. W., Leech, N. L., & Morgan, G. A. (2006). *SPSS for introductory statistics: Use and interpretation*. New York, NY: Psychology Press – Taylor & Francis Group.
- Bass, F. & Dizon, N.Z. (2006, April 18). States help schools dodge No Child: Accused of “gaming the system.” *Chicago Sun Times*.
- Berg, B. L. (2004). *Qualitative research methods for the social sciences*. Boston, MA: Pearson Allyn and Bacon.
- Berliner, D. C. (2009). The incompatibility of high-stakes testing and the development of Skills for the 21st century. In R. Marzano (ed.), *Anthology, 2009*. Bloomington, IN: Solution Tree Press.
- Booher-Jennings, J. (2006, June). Rationing education in an era of accountability. *Phi Delta Kappan*, 87. Retrieved August 12, 2009 from http://www.pdkintl.org/kappan/k_v87/k0606toc.htm
- Callahan, D. (2004). *The cheating culture: Why more Americans are doing wrong to get ahead*. Orlando, FL: Harcourt, Inc.
- Cimbricz, S. (2002). State-mandated testing and teachers' beliefs and practice. *Education Policy Analysis Archives*, 10(2). Retrieved August 12, 2009 from <http://epaa.asu.edu/epaa/v10n2.html>
- Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating. Quantitative and qualitative research*. (3rd ed.). Upper Saddle River, NJ: Merrill Prentice Hall.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Day, C., Sammons, P., & Gu, Q. (2008). Combining qualitative and quantitative methodologies in research on teachers' lives, work, and effectiveness: From integration to synergy. *Educational Researcher*, 37(6), 330-342.
- Diamantopoulus, A. & Schlegelmilch, B. B. (2006). *Taking the fear out of data analysis: A step-by-step approach*. Bedford Row, London: Thompson Learning.
- Dorn, S. (1998). The political legacy of school accountability systems. *Education Policy Analysis Archives*, 6(1). Retrieved March 15, 2010 from <http://olam.ed.asu.edu/epaa/v6n1.html>
- Ercikan, K., & Wolff-Michael, R. (Eds.) (2009). *Generalizing from educational research: Beyond qualitative and quantitative polarization*. New York: Routledge.
- Fausset, R. (2010, February 17). Georgia investigates abnormal test scores. *Los Angeles Times*. Retrieved March 9, 2010 from <http://articles.latimes.com/2010/feb/17/nation/la-na-georgia-scores17-2010feb17>
- Fessenden, F. (2007a, September 9). Schools under scrutiny over cheating. *The New York Times*. Retrieved August 15, 2009 from <http://www.nytimes.com/2007/09/09/nyregion/nyregionspecial2/09Rcheating.html>
- Fessenden, F. (2007b, June 24). Uniondale math results voided as cheating is cited. *The New York Times*. Retrieved August 15, 2009 from <http://www.nytimes.com/2007/06/24/nyregion/nyregionspecial2/24topiccli.html>

- Firestone, W. A., Camilli, G., Yurecko, M., Monfils, L., & Mayrowetz, D. (2000). State standards, socio-fiscal context and opportunity to learn in New Jersey. *Education Policy Analysis Archives*, 8(35). Retrieved August 12, 2009 from <http://olam.ed.asu.edu/epaa/v8n35/>
- Flowers, A. B. (2010). Blazing an evaluation pathway: Lessons learned from applying utilization-focused evaluation to a conservation education program. *Evaluation and Program Planning*, 33(2), 165-171.
- Gootman, E. (2008, December 3). School official to be fired after cheating is found. *The New York Times*. Retrieved August 15, 2009 from <http://www.nytimes.com/2008/12/04/education/04cheat.html>
- Greene, J. C., Caracelli, V. J., & Graham, W. D. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11(3), 255-274.
- Haladyna, T. M. (2002). *Essentials of standardized achievement testing: Validity and accountability*. Boston, MA: Allyn and Bacon.
- Haladyna, T., Nolen, S. B., & Haas, N. S. (1991). Raising standardized test scores and the origins of test score pollution. *Educational Researcher*, 20(5), 2-7.
- Haney, W. (2000). The myth of the Texas miracle in education. *Education Analysis Policy Archives*, 8(41) Retrieved August 12, 2009 from <http://epaa.asu.edu/epaa/v8n41>
- Haver, J. (2004, September 13). Cheating on AIMS tests: It's not students doing it. *Arizona Republic*.
- Heubert, J. P. & Hauser, R. M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press. Retrieved August 12, 2009 from at <http://www.nap.edu/html/highstakes>
- Johnson, D. D., Johnson, B., Farenga, S. J., & Ness, D. (2007). *Stop high-stakes testing: An appeal to America's conscience*. Lanham, MD: Rowman & Littlefield.
- Johnson, R. B. & Onwuegbuzie, A. J. (2004) Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Jones, M. G, Jones, B. D., & Hargrove, T. (Eds.). (2003). *The unintended consequences of high-stakes testing*. Lanham, MD: Rowman & Littlefield Publishers, Inc.
- Kelleher, M. (1999, June). Dropout rate climbs as schools dump truants. *Catalyst*. Retrieved from <http://www.catalyst-chicago.org/arch/06-99/069pushouts.htm>
- Koretz, D. (1996). Using student assessments for educational accountability. In E.A. Hanushek & D.W. Jorgenson (Eds.) *Improving America's schools: The role of incentives*. Washington, D.C.: National Academy Press.
- Kreitzer, A.E., Madaus, G.F., & Haney, W. (1989). Competency testing and dropouts. In L. Weis, E. Farrar, & H.G. Petrie (Eds.) *Dropouts from school: Issues, dilemmas, and solutions*. Albany, NY: State University of New York Press.
- Linn, R. L. (2000). Assessments and accountability. *Education Researcher*, 29(2), 4-15.
- Linn, R. L., Graue, M. E., & Sanders, N. M. (1990). Comparing state and district results to national norms: The validity of the claims that "everyone is above average." *Educational Measurement: Issues and Practice*, 9(3), 5-14.
- McGill-Franzen, A. & Allington, R. L. (1993). Flunk'em or get them classified: The contamination of primary grade accountability data. *Educational Researcher*, 22(1), 19-22.
- McNeil, L. (2000). *Contradictions of school reform*. New York, NY: Routledge.
- Madaus, G. F. & Clarke, M. (2001). The adverse impact of high stakes testing on minority students: Evidence from one hundred years of test data. In G. Orfield & M. Kornhaber (Eds.). *Raising standards or raising barriers? Inequality and high-stakes testing in public education*. New York, NY: The Century Foundation Press.

- Madaus, G. F., West, M. M., Harmon, M. C., Lomax, R. G., & Viator, K. A. (1992). *The influence of testing on teaching math and science in grades 4–12*. Chestnut Hill, MA: Center of Study of Testing, Evaluation, and Educational Policy, Boston College.
- Market Tools, Inc. (2008) *Zoomerang Online Surveys*. www.zoomerang.com
- May, M. (2000, October 4). State fears cheating by teachers: 51 schools left off cash award list. *The San Francisco Chronicle*. Retrieved August 12, 2009 from <http://www.sfgate.com/cgi-bin/article.cgi?file=/chronicle/archive/2000/10/04/MN100403.DTL>
- Mehrens, W. A. (1998). Consequences of assessment: What is the evidence? *Education Policy Analysis Archives*, 6(13). Retrieved August 12, 2009 from <http://olam.ed.asu.edu/epaa/v6n13.html>
- Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Miretzky, D. (2009). A view of research from practice: Voices of teachers. *Theory Into Practice*, 46(4), 272–280.
- Nichols, S., & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts America's schools*. Cambridge, MA: Harvard Education Press.
- Nichols, S. L., Glass, G. V, & Berliner, D. C. (2005). *High-stakes testing and student achievement: Problems for the No Child Left Behind Act*. Education Policies Laboratory. Retrieved February 27, 2010 from <http://www.asu.edu/educ/eps/EPRU/documents/EPSE-0509-105-EPSE.pdf>
- Nichols, S. L., Glass, G. V, & Berliner, D. C. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning? *Education Policy Analysis Archives*, 14(1). Retrieved August 12, 2009 from <http://epaa.asu.edu/epaa/v14n1>
- No Child Left Behind Act. (2001). *The No Child Left Behind Act*. Retrieved August 12, 2009 from <http://www.ed.gov>
- Ohanian, S. (2010, April 7). *Race to the top: Delaware's application for initial funding*. Retrieved April 10, 2010 from http://susanoonian.org/show_commentary.php?id=794
- Olson, L. (2000, July 12). High-stakes tests jeopardizing Hispanics, panel warns. *Education Week*. Retrieved August 12, 2009 from <http://www.edweek.org/ew/index.html>
- Orfield, G. & Kornhaber, M. L. (Eds.). (2001). *Raising standards or raising barriers? Inequality and high-stakes testing in public education*. New York, NY: The Century Foundation Press.
- Phelps, R. (2005). *The war on standardized testing: Kill the messenger*. New Brunswick, NJ: Transaction Publishers.
- Popham, W. J. (2007). *Classroom assessment: What teachers need to know* (5th ed.). Needham Heights, MA: Allyn & Bacon.
- Radin, B. (2000). *Beyond Machiavelli: Policy analysis comes of age*. Washington, DC: Georgetown University Press.
- Ross, J. 2005. Arts education and the newer public good. *Arts Education Policy Review*, 106(3), 3-7.
- Sacks, P. (1999). *Standardized minds: The high price of America's testing culture and what we can do to change it*. Cambridge, MA: Perseus Publishing.
- Schrag, P. (2000, August). High stakes are for tomatoes. *The Atlantic Monthly*. Retrieved August 12, 2009 from <http://www.theatlantic.com/issues/2000/08/schrag.htm>
- Sheldon, K. M. & B. J. Biddle (1998). Standards accountability and school reform: Perils and pitfalls. *Teachers College Record*, 100(1), 164–180.
- Shepard, L.A. (1990). Inflated test score gains: Is the problem old norms or teaching to the test? *Educational Measurement: Issues and Practice*, 9(3), 15–22.
- Smith, M.L. & Fey, P. (2000). Validity and accountability in high-stakes testing. *Journal of Teacher Education*, 51(5), 334-344.

- Sociology of Education Research Group. (1998, December). *Evaluation of academic performance in the Houston Independent School District*. Houston, TX: Center for Houston's Future.
- Sun, W. Chou, C. P., Stacy, A. W., Ma, H., Unger, J., & Gallaher, P. (2007, February). SAS and SPSS macros to calculate standardized Cronbach's alpha using the upper bound of the phi coefficient for dichotomous items. *Behavior Research Methods*, 39(1), 71-81.
- Swope, K. & Miner, B. (2000). A vision of school reform. In K. Swope & B. Miner (Eds.) *Failing our kids: Why the testing craze won't fix our schools*. Milwaukee, WI: Rethinking Schools, Ltd.
- Thompson, S. (2001). The authentic standards movement and its evil twin. *Phi Delta Kappan*, (82)5, 358-362.
- Urdu, T. C., Paris, S. G. (1994). Teachers' perceptions of standardized achievement tests. *Educational Policy*, 8(2), 137-157.

Appendix

Survey Instrument

Section 1 - Demographics

1. Have you ever been a teacher in the public school system? (Yes, No)
2. How many years of teaching experience do you have? (None, 1-5, 6-10, 11-15, 16-20, 21-25, 26+)
3. What is your primary role in your current position?
4. What is your highest degree earned? (BA, BS, MA, Ed.D, Ph.D)
5. In what type of school do you currently work? (Charter, Private, Public)
6. In what type of area is your school located? (Rural, Suburban, Urban)
7. Is the school in which you currently teach a Title 1 school?
8. What label has your school received from the Arizona Department of Education? (Excelling, Highly performing, Performing plus, Performing, Under performing, Failing, Do not know)

Section 2 – Awareness of Others’ Behaviors

9. Do you know of another educator who has... a. Erased and changed test answers; b. Gave answers to students; c. Pointed to correct answers; d. Changed student identification numbers in order to eliminate low-scoring students’ results; e. Not tested academically weak students; f. Asked students to be absent during testing days; g. Encouraged others to cheat; h. Encouraged students to redo problems; i. Coached students during the test; j. Created cheat sheets for students to help prepare them for tests; k. Read questions to students when not allowed to do so; l. Gave students extra time on tests; m. Allowed students to work on tests during lunch, recess, or specials areas; n. Allowed students to figure out answers on scratch paper for them to check prior to having students mark answer sheets; o. Made copies of tests in order to teach to the tests; p. Wrote down questions to prepare for following years’ tests; q. Wrote down vocabulary words to teach to following years’ classes; r. Kept test booklets in order to familiarize students with the tests; s. Left materials on the walls knowing it would help students.
10. Please elaborate on these or other similar incidences here.

Section 3 – Self-Reported Behaviors

11. How many times have you... a. Erased and changed test answers; b. Gave answers to students; c. Pointed to correct answers; d. Changed student identification numbers in order to eliminate low-scoring students’ results; e. Not tested academically weak students; f. Asked students to be absent during testing days; g. Encouraged others to cheat; h. Encouraged students to redo problems; i. Coached students during the test; j. Created cheat sheets for students to help prepare them for tests; k. Read questions to students when not allowed to do so; l. Gave students extra time on tests; m. Allowed students to work on tests during lunch, recess, or specials areas; n. Allowed students to figure out answers on scratch paper for them to check prior to having students mark answer sheets; o. Made copies of tests in order to teach to the tests; p. Wrote down questions to prepare for following years’ tests; q. Wrote down vocabulary words to teach to following years’ classes; r. Kept test booklets in order to familiarize students with the tests; s. Left materials on the walls knowing it would help students.
12. Please elaborate on these or other similar incidences here.

About the Authors

Audrey Amrein-Beardsley
Arizona State University

David C. Berliner
Arizona State University

Sharon Rideau
Arizona State University

Email: audrey.beardsley@asu.edu

Dr. Audrey Amrein-Beardsley is currently an Associate Professor in the Mary Lou Fulton Teachers College. Her research interests include educational policy, research methods, and high-stakes tests, assessments and measurement. She also researches aspects of teacher quality and teacher education.

David C. Berliner is Regents' Professor Emeritus at Arizona State University. His recent research is on high-stakes testing and also on how poverty affects school achievement. His work has also focused on the study of teaching, teacher education, and the profession of teaching.

Sharon Rideau graduated with her Ed.D. from Arizona State University in 2009. She is currently teaching in California and plans to continue teaching while continuing to research cheating by educators on standardized achievement tests.

education policy analysis archives

Volume 18 Number 14

June 25th 2010

ISSN 1068-2341



Readers are free to copy, display, and distribute this article, as long as the work is attributed to the author(s) and **Education Policy Analysis Archives**, it is distributed for non-commercial purposes only, and no alteration or transformation is made in the work. More details of this Creative Commons license are available at <http://creativecommons.org/licenses/by-nc-sa/3.0/>. All other uses must be approved by the author(s) or **EPAA**. **EPAA** is published by the Mary Lou Fulton Institute and Graduate School of Education at Arizona State University. Articles are indexed EBSCO Education Research Complete, [Directory of Open Access Journals](#), ERIC, H.W. WILSON & Co, QUALIS – A 2 (CAPES, Brazil), SCOPUS, SOCOLAR-China.

Please contribute commentaries at <http://epaa.info/wordpress/> and send errata notes to Gustavo E. Fischman fischman@asu.edu

education policy analysis archives
editorial board

Editor **Gustavo E. Fischman** (Arizona State University)

Associate Editors: **David R. Garcia & Jeanne M. Powers** (Arizona State University)

Jessica Allen University of Colorado, Boulder

Gary Anderson New York University

Michael W. Apple University of Wisconsin,
Madison

Angela Arzubiaga Arizona State University

David C. Berliner Arizona State University

Robert Bickel Marshall University

Henry Braun Boston College

Eric Camburn University of Wisconsin, Madison

Wendy C. Chi* University of Colorado, Boulder

Casey Cobb University of Connecticut

Arnold Danzig Arizona State University

Antonia Darder University of Illinois, Urbana-
Champaign

Linda Darling-Hammond Stanford University

Chad d'Entremont Strategies for Children

John Diamond Harvard University

Tara Donahue Learning Point Associates

Sherman Dorn University of South Florida

Christopher Joseph Frey Bowling Green State
University

Melissa Lynn Freeman* Adams State College

Amy Garrett Dikkers University of Minnesota

Gene V Glass Arizona State University

Ronald Glass University of California, Santa Cruz

Harvey Goldstein Bristol University

Jacob P. K. Gross Indiana University

Eric M. Haas WestEd

Kimberly Joy Howard* University of Southern
California

Aimee Howley Ohio University

Craig Howley Ohio University

Steve Klees University of Maryland

Jaekyung Lee SUNY Buffalo

Christopher Lubienski University of Illinois,
Urbana-Champaign

Sarah Lubienski University of Illinois, Urbana-
Champaign

Samuel R. Lucas University of California,
Berkeley

Maria Martinez-Coslo University of Texas,
Arlington

William Mathis University of Colorado, Boulder

Tristan McCowan Institute of Education, London

Heinrich Mintrop University of California,
Berkeley

Michele S. Moses University of Colorado, Boulder

Julianne Moss University of Melbourne

Sharon Nichols University of Texas, San Antonio

Noga O'Connor University of Iowa

João Paraskveva University of Massachusetts,
Dartmouth

Laurence Parker University of Illinois, Urbana-
Champaign

Susan L. Robertson Bristol University

John Rogers University of California, Los Angeles

A. G. Rud Purdue University

Felicia C. Sanders The Pennsylvania State
University

Janelle Scott University of California, Berkeley

Kimberly Scott Arizona State University

Dorothy Shipps Baruch College/CUNY

Maria Teresa Tatto Michigan State University

Larisa Warhol University of Connecticut

Cally Waite Social Science Research Council

John Weathers University of Colorado, Colorado
Springs

Kevin Welner University of Colorado, Boulder

Ed Wiley University of Colorado, Boulder

Terrence G. Wiley Arizona State University

John Willinsky Stanford University

Kyo Yamashiro University of California, Los Angeles

* Members of the New Scholars Board

archivos analíticos de políticas educativas
consejo editorial

Editor: **Gustavo E. Fischman** (Arizona State University)

Editores. Asociados **Alejandro Canales** (UNAM) y **Jesús Romero Morante** (Universidad de Cantabria)

- Armando Alcántara Santuario** Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México
- Claudio Almonacid** Universidad Metropolitana de Ciencias de la Educación, Chile
- Pilar Arnaiz Sánchez** Universidad de Murcia, España
- Xavier Besalú Costa** Universitat de Girona, España
- Jose Joaquin Brunner** Universidad Diego Portales, Chile
- Damián Canales Sánchez** Instituto Nacional para la Evaluación de la Educación, México
- María Caridad García** Universidad Católica del Norte, Chile
- Raimundo Cuesta Fernández** IES Fray Luis de León, España
- Marco Antonio Delgado Fuentes** Universidad Iberoamericana, México
- Inés Dussel** FLACSO, Argentina
- Rafael Feito Alonso** Universidad Complutense de Madrid, España
- Pedro Flores Crespo** Universidad Iberoamericana, México
- Verónica García Martínez** Universidad Juárez Autónoma de Tabasco, México
- Francisco F. García Pérez** Universidad de Sevilla, España
- Edna Luna Serrano** Universidad Autónoma de Baja California, México
- Alma Maldonado** Departamento de Investigaciones Educativas, Centro de Investigación y de Estudios Avanzados, México
- Alejandro Márquez Jiménez** Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México
- José Felipe Martínez Fernández** University of California Los Angeles, USA
- Fanni Muñoz** Pontificia Universidad Católica de Perú
- Imanol Ordorika** Instituto de Investigaciones Economicas – UNAM, México
- Maria Cristina Parra Sandoval** Universidad de Zulia, Venezuela
- Miguel A. Pereyra** Universidad de Granada, España
- Monica Pini** Universidad Nacional de San Martín, Argentina
- Paula Razquin** UNESCO, Francia
- Ignacio Rivas Flores** Universidad de Málaga, España
- Daniel Schugurensky** Universidad de Toronto-Ontario Institute of Studies in Education, Canadá
- Orlando Pulido Chaves** Universidad Pedagógica Nacional, Colombia
- José Gregorio Rodríguez** Universidad Nacional de Colombia
- Miriam Rodríguez Vargas** Universidad Autónoma de Tamaulipas, México
- Mario Rueda Beltrán** Instituto de Investigaciones sobre la Universidad y la Educación, UNAM México
- José Luis San Fabián Maroto** Universidad de Oviedo, España
- Yengny Marisol Silva Laya** Universidad Iberoamericana, México
- Aida Terrón Bañuelos** Universidad de Oviedo, España
- Jurjo Torres Santomé** Universidad de la Coruña, España
- Antoni Verger Planells** University of Amsterdam, Holanda
- Mario Yapu** Universidad Para la Investigación Estratégica, Bolivia

arquivos analíticos de políticas educativas
conselho editorial

Editor: **Gustavo E. Fischman** (Arizona State University)
Editores Associados: **Rosa Maria Bueno Fisher** e **Luis A. Gandin**
(Universidade Federal do Rio Grande do Sul)

- | | |
|--|---|
| Dalila Andrade de Oliveira Universidade Federal de Minas Gerais, Brasil | Jefferson Mainardes Universidade Estadual de Ponta Grossa, Brasil |
| Paulo Carrano Universidade Federal Fluminense, Brasil | Luciano Mendes de Faria Filho Universidade Federal de Minas Gerais, Brasil |
| Alicia Maria Catalano de Bonamino Pontifícia Universidade Católica-Rio, Brasil | Lia Raquel Moreira Oliveira Universidade do Minho, Portugal |
| Fabiana de Amorim Marcello Universidade Luterana do Brasil, Canoas, Brasil | Belmira Oliveira Bueno Universidade de São Paulo, Brasil |
| Alexandre Fernandez Vaz Universidade Federal de Santa Catarina, Brasil | Antônio Teodoro Universidade Lusófona, Portugal |
| Gaudêncio Frigotto Universidade do Estado do Rio de Janeiro, Brasil | Pia L. Wong California State University Sacramento, U.S.A |
| Alfredo M Gomes Universidade Federal de Pernambuco, Brasil | Sandra Regina Sales Universidade Federal Rural do Rio de Janeiro, Brasil |
| Petronilha Beatriz Gonçalves e Silva Universidade Federal de São Carlos, Brasil | Elba Siqueira Sá Barreto Fundação Carlos Chagas , Brasil |
| Nadja Herman Pontifícia Universidade Católica – Rio Grande do Sul, Brasil | Manuela Terrasêca Universidade do Porto, Portugal |
| José Machado Pais Instituto de Ciências Sociais da Universidade de Lisboa, Portugal | Robert Verhine Universidade Federal da Bahia, Brasil |
| Wenceslao Machado de Oliveira Jr. Universidade Estadual de Campinas, Brasil | Antônio A. S. Zuin Universidade Federal de São Carlos, Brasil |