

## FROM THE PRESIDENT

*Wim J. van der Linden, CTB/McGraw-Hill*



When you read this message you're probably preparing your visit to Philadelphia for our annual meeting. I hope you're also planning to attend the NCME Breakfast—one of the best traditions of our organization.

A major announcement at the Business Meeting will be a survey of the NCME membership regarding the future of our annual meetings. Right now, NCME is close to an extension of its current agreement for meeting services with AERA for 2015-2016. NCME has always cherished its relationship with AERA, and we are happy the agreement will enable us to join AERA in the celebration of its centennial in Washington, DC, in 2016. During the 2016 AERA Opening Plenary, Presidential Address, and Awards Luncheon, NCME will not schedule any sessions or events to allow its members to attend these highlights.

Needless to say we have always been happy to rely on AERA for its meeting services. On the other hand, both annual meetings have grown considerably in size. Consequently, our venues, although in the same city, have consistently been too far apart to allow the attendants to switch between sessions for several years now. And, as all past program co-chairs and presidents have experienced, coordinating the preparations for the two annual meetings has become difficult. For instance, our meeting program is usually finalized by the end of November, but we are unable to publish a program until a few weeks before the meeting due to conflict resolution with AERA. Besides, during our communications with AERA it has become clear that further extension of our agreement is no longer something that will happen automatically.

As annual meeting planning has to begin many years in advance, a while ago an ad hoc NCME committee (David Frisbie, Chair; Terry Ackerman; and Anne Fitzpatrick) was charged with determining the feasibility of planning and conducting our annual meetings without an agreement with AERA. This would primarily mean that NCME takes full responsibility for the organization of its own meeting, using its own resources (central office; committees) to make hotel choices, support meeting registration, market the meeting, etc. The choice would allow NCME to meet at a different time or place but certainly does not imply that.

With the help of our central office, the committee has produced an extremely informative report, exploring all pros and cons of every possible scenario. Whatever happens to or will be decided on the future of its annual meeting, NCME now has a comprehensive inventory of all possible consequences and will be fully prepared to act. Clearly, the Board will need input from our membership and continued communication with AERA before it is able to make any decision as to the annual meetings beyond 2016.

The committee has also pointed at an obvious omission in our current governance structure for the annual meeting and recommended the establishment of a standing annual meeting committee. Right now, NCME lacks continuity in basic planning aspects of its annual meetings, and communication between the groups that implement the various aspects of the meeting is not always optimal. A standing committee would not only ensure the continuity and enhance the communication but also provide permanent oversight for duties delegated to AERA and TRG, our current contractors for annual meeting services. In its recent meeting, the Board unanimously accepted this recommendation, and has found Terry Ackerman willing to chair the committee. The committee will also be in charge of the upcoming survey among the NCME membership about the future of the annual meeting. Terry will introduce his committee at the Business Meeting in Philadelphia and tell you more about the survey.

So, if you walk around in our conference hotel in Philadelphia do not only enjoy the fascinating program Paul De Boeck and Kathleen Scalise have prepared but also reflect on the enormous amount of organizational work by your program chairs, committees, central office, and board that goes into the preparation of the meeting. And if we happen to run into one another, don't hesitate sharing with me any questions or comments you may have regarding its future.

See you in Philadelphia!

---

## GREETINGS FROM THE EDITOR

*Susan Davis-Becker, Alpine Testing Solutions*

Happy **SPRING** to everyone! We have some interesting reads in this issue for all members of NCME! The issue kicks off with the final presidential column from Dr. Wim van der Linden – please join me in congratulating Wim on a successful presidential term. Our 2014 graduate student columnist is Diane Talley from the University of North Carolina, Chapel Hill (welcome Diane!) and she provides some great insight on the broad world of psychometrics. Our Spotlight member in this issue is Dr. Laurie Wise – the next president of NCME! As a special feature in this issue, we had a few members provide perspectives on a very important topic in our field – cheating. Specifically, we look at how “accidental” (unintentional) cheating occurs in different types of testing programs. This is followed by an update from the NCME *Standards and Test Use Committee*. We have also included lots of important information about the upcoming annual meeting as well as other events going on in the field. Lastly, we conclude with thoughts from Alan Nicewander on the life and work of Howard Mitzel who passed away in January.

Looking forward to seeing everyone next month in Philly!

---

## GRADUATE STUDENT CORNER:

### WHEN I GROW UP, I WANT TO BE A PSYCHOMETRICIAN

*Diane Talley, University of North Carolina, Chapel Hill*

Is this what you said when you were a child? Few people know what a psychometrician is until they stumble upon the profession at some point in college or in the early stages of a career. If you read the Spotlight column in these newsletters, they all begin with a similar story of an accidental entry into the profession. I suspect your story is similar.

What does a psychometrician actually do? There is so much variety in the field that it is impossible to distill it into a single job role. We know, generally speaking, that psychometricians use measurement and research theory to design, implement, and evaluate assessment programs and instruments, but the types of programs, constructs and content measured, subject matter experts involved in the process, relevant testing policy, and the specific role the psychometrician plays may vary significantly. Psychometricians play key roles in a number of disciplines, including education, traditional and Industrial-Organizational (I/O) psychology, credentialing, program evaluation, academia and research. Newer fields have emerged that cross disciplines such as assessment technology and security.

The breadth of roles is critical for graduate students to be aware of, yet too often is overlooked. Students who focus on a single aspect of psychometrics while in their graduate programs may unnecessarily limit their career options down the road. For example, those in a measurement program in education may not be aware of psychometric roles in credentialing.

Psychometric positions in credentialing have much to offer graduates searching for interesting and challenging experiences in a very diverse industry. Credentialing is a broad term that covers the assessment of persons against specific job-related requirements and encompasses both professional certification and licensure. I focus here on certification programs and the unique experiences they can offer graduates.

The measurement theory and methodology across disciplines is informed by the same body of research, but the application may vary considerably. One such difference is the potential diversity of the tested subject areas. Certification spans a multitude of industries including medical, technology, government, and financial for example. Psychometricians working with certification programs may be involved in the overall design of testing programs including establishment of testing policies in addition to the design of the examinations themselves.



An interesting aspect of this field is the interaction between certification programs, the law (especially personnel laws), and accreditation. Many certification programs are accredited, meaning that they obtain a third party verification of quality and adherence to established criteria from organizations such as ANSI and ICE. The criteria are based on the Standards for Educational and Psychological Measurement (AERA, APA, NCME, 1999), with which you are hopefully becoming intimately familiar through your graduate studies. Assisting programs in the accreditation process brings in an element of program evaluation, assessing how all of the elements of a testing system compare against the established standards.

The measurement models and methodology you are currently studying in graduate school such as IRT, Classical Test Theory, equating, and standard setting are all necessary for a career working with professional certification programs. An additional skill you may not be familiar with if you are focusing on educational measurement or research methods is conducting role delineation studies. These are the studies that define what will be tested on a professional certification examination. While training specifically for these types of studies is typically found in I/O Psychology programs, study of general measurement theory, validity theory (particularly content-related validity evidence), and survey methodology will provide a strong foundation for conducting such studies.

This description of psychometric work in certification is a generalization, of course, and has its share of caveats. The role of a psychometrician in any field can depend in large part on the size and type of the testing organization in which they work. The organization of its psychometric department can have a similar impact.

To give my fellow graduate students a better look inside the life of a psychometrician who has worked in certification, I asked Dr. Stephen Johnson who has worked across education, certification, and program evaluation, to share some of his experience and insight. I started by asking him to tell us a little about his background and career choices.

“I initially trained as a psychologist, with the intention of becoming anything BUT a clinical psychologist. I attended what we called a ‘rats and stats’ school of psychology – lots of experimental psychology, analysis work, and escaped rats. I went to work for an Australian federal government agency that eventually involved a lot of training and interviewing, and data analysis. After working for a few years I went back to school to complete a teaching diploma as I felt that it would be a useful tool to help with training and presentations. While at school I had long discussions with faculty and fixed my mind on a PhD and research career, mostly in the education world. I spent 12 years working in educational assessment and evaluation for K-university programs before taking a position in a privately held company that supported the development of adult credentialing programs.”

Since Dr. Johnson has worked across assessment fields, I asked him to comment on how he views differences in the role of a psychometrician in certification versus education. He indicated that that from a technical aspect they are the same. He states that “a client has an interest in establishing some understanding of a psychological trait and wants to develop an assessment protocol that meets: a. legal criteria, b. costs and resource capabilities of the organization, c. enables effective communication with a wide variety of stakeholders.”

An interesting point Dr. Johnson made was with regards to stakeholders related to certification programs, which are very different from education. “With educational assessment, parents and politicians are significant stakeholders, whilst the actual students/examinees have little voice in the process.” In certification, the examinees are adults and frequently have voice in the process either through volunteer involvement with the program or by expressing opinions to program staff and the industry at large.

One of the interesting features of working with certification programs is the different subject matter one works with. In education, they are typically teachers or policy makers. In certification they are experts from whichever industry the program is situated. One day you may be doing a standard setting with a group of electricians, then the next a content validity study for data analysts, and on another working with a nursing group to set testing policy. As Dr. Johnson put it, “Clearly from my history, I like variety, which is achievable by folks in the field.”

In addition to sharing his experiences as a psychometrician, Dr. Johnson gave some sage advice about what experiences he sees most important in preparing for a career as a psychometrician.

“I used to think it was the networking opportunities, but in a large, international, and somewhat fragmented field (e.g., educational assessment rarely mix with I/O or vice versa), I am no longer clear it is the most beneficial. Conversely, an argument could be made that relying on a supervisor for your next connections and steps may be limiting the scope of the type of work you could get involved in. Going to one or two large conferences (AERA, APA, SIOP) definitely helps broaden the concept of what is out there.

I would always recommend teaching - done well it helps sharpen your thinking and works to prepare you for talking with a diverse group of folks - VERY helpful when working with stakeholder groups.

See if you can work on other projects, especially ones that push you out of your comfort zone. It is too easy to do repeats of the same work (e.g., analyze this data set ... now this one), which may help pay the bills but won't help you understand assessment design, delivery and analysis."

Hopefully this has given you something to think about as you progress through your graduate programs and consider your ideal career. For those interested in more information about psychometrics in credentialing you may want to review the accreditation organizations discussed earlier (ANSI and ICE) and the AERA SIG for Professional Certification and Licensure. I would also recommend looking through various job descriptions both on job boards and on testing company websites. Many of the middle to large-sized testing companies have divisions that focus on credentialing as well as education.

For more information:

ANSI: <http://ansi.org/>

ICE: <http://www.credentialingexcellence.org/>

AERA SIG for Professional Certification and Licensure:

<http://www.aera.net/SIG080/ProfessionalLicensureandCertificationSIG80/tabid/11605/Default.aspx>

### References:

American Educational Research Association, American Psychological Association, & National Council of Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: AERA.

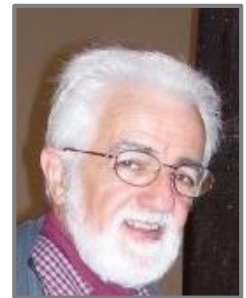
---

## SPOTLIGHT ON THE PEOPLE WHO MAKE OUR ORGANIZATION GREAT

*This month our spotlight is on Laress (Laurie) Wise. In addition to being a Principal Scientist at Human Resources Research Organization (HumRRO), Laurie is the incoming president of NCME*

### How did you get into the field?

My undergraduate major was in mathematics and then I added psychology, because it was the 60's and mathematics seemed very dry and irrelevant to the human condition. I went to graduate school in psychology, thinking I could work on mathematical models for attitude change (e.g., Vietnam War) and related psychological constructs. I discovered that mathematical models were of little use without solid measures to use in building and testing the models, and so drifted into measurement. We had a graduate seminar using a fascinating new book by Lord and Novick that included an introduction to item response theory. From graduate school, I went to work at AIR on a career counseling project, but soon got pulled into work on a graduate school admissions program to resolve an equating problem. It turns out that you cannot assume that items at the end of the test behave in the same way when they are placed earlier in the test. I've been a psychometrician ever since.



### If you weren't in this field, what would you do?

I really have no idea. As a youth, I had some interest in working on the space program. That might have worked out. I sometimes think about going into politics, because the current bar for success appears to be quite low.

### What advice would you have for graduate students who want to get into this field?

A solid background in statistics is essential for both measurement work and research. Beyond that there are emerging areas of research, such as cognitive models for diagnostic testing, multistage-multivariate testing, and nailing down types of evidence in Evidence-Centered Design. My advice would be to find a good professor and ask him or her a lot of questions; go to meetings and listen to a wide variety of presentations to see what interests you; and, of course, read the journals. Finally, when you find topics of interest, commit to presentations and publications on your work on the topic.

### What do you like to do for fun outside of work?

Right now, there is not a lot beyond work. I do like to grow vegetables (they don't talk back much and are usually good to eat), hike, and babysit my adorable granddaughters. I hope to have more time for outside activities when my NCME presidential year concludes.

**What would you say has been one of the biggest innovations in psychometrics in the last decade or two?**

Evidence-Centered Design.

**When you go to conferences, how do you pick what sessions to attend?**

I used to pick on the basis of topics of greatest interest. Now, I mostly go to listen to friends talk about their work, which also tends to be on topics of interest.

**Who has been a significant influence in your professional life?**

I'd have to say Donald McLaughlin. He was one of my professors at Berkeley, using protocol analyses to generate mathematical models of cognition and behavior. He also hired me to work at AIR and supervised my early work. He has always been a creative, out-of-the-box thinker and I have tried to follow his example.

---

## A LOOK AT UNINTENTIONAL CHEATING IN THE TESTING INDUSTRY

Any reviews of recent literature, conference publications, or testing news in the media will clearly show that one of the most talked about topics in testing right now is cheating. Numerous scholars in our field are providing us with valuable information on ways to detect and prevent cheating on high-stakes examinations. In this newsletter, we decided to take a look at one specific type of cheating – unintentional (accidental). Specifically, we were wondering if there are instances in which examinees are obtaining inappropriate information or help for a high-stakes exam without realizing they are violating testing policies. We asked colleagues from different parts of the testing industry to provide their perspectives on this topic.

### Professional Credentialing

*Ardeshir Geranpayeh, Cambridge English Language Assessment, UK*

Test and data integrity has always been an important aspect of any credentialing testing agency. Cheating in tests is not a new phenomenon and has existed as long as exams have been used for decision making. One could trace it back over a thousand years to Chinese civil service exams. The NCME has always been in the forefront of combating cheating in exams and providing guidelines to address such issues. There are at least 5 explicit statements in the *Standards for Educational and Psychological Testing* (AERA, APA, and NCME, 1999) for the prevention of cheating: Standards 8.7, 11.7, 13.10, 13.11, and 15.9. In recent years, due to various media coverage, the topic of cheating has been widely debated in various academic and public fora. The NCME has taken a lead on this and has produced a document on Test and Data Integrity (2012), allocated two Presidential invited sessions at the 2013 annual meeting and at least one at the 2014 meeting in addition to accepting various individual papers on this topic. The topic of cheating and test security in general has been extensively discussed in various publications. Cizek (1999), Wollack and Fremer (2013), and Geranpayeh (2014) are just a few examples.



Credentialing testing programs have a duty to protect the abuse of their test results by fraudulent means. Most, if not all, cheating detection techniques designed for this purpose are based on the assumption that the test taker or other people involved in the inappropriate use of test results are deceptively and deliberately cheating in the exams in order to secure unfair or unlawful gain. Therefore, once the cheating is detected the blame is allocated to the test taker or other people believed to be involved in that process. What is often forgotten is that the cheating detection techniques can only identify individuals who have suspicious results. What they cannot do is to say whether the suspicious act was deliberate. The evidence for the latter has to come from other sources available to the testing organization and is subject to legal challenges in some jurisdiction.

This column distinguishes Plagiarism from Cheating. Geranpayeh (2014) defines plagiarism as “a special act of cheating associated with essay writing” (p. 981). Although plagiarism is a serious offence in academic contexts, its nebulous boundary with copying (legitimate) is not always clear-cut. Most reports about students frequently cheating at colleges and universities refer to plagiarism. Many higher educational institutions now have guidelines for students to avoid plagiarism and cheating. Whether plagiarism (cheating) is intentional in such contexts is out of the scope of this paper. See Geranpayeh (2014) for a short review on plagiarism.

### **What is unintentional (accidental) cheating?**

Cizek (1999) defines cheating as “any action that violates the rules for administering a test”. Furthermore, Cizek (2001, p. 5) expands on this by describing cheating as “any behavior that gives an examinee an unfair advantage over other examinees, or any action on the part of an examinee or test administrator that decreases the accuracy of the intended inferences arising from the examinee’s test score or performance”. This broad definition does not differentiate between intentional and unintentional cheating. If some candidates get unfair advantage over others by breaking test administration rules, little matters with respect to their intention.

One may argue that there must be a difference when test administration rules are broken unintentionally. There is no clear definition for unintentional cheating. Grijalva et al. (2006) associate unintentional cheating with an act of unplanned cheating where students attempt to cheat in a moment of panic when they realize that they have no idea how to answer a question. In such cases the students often come up with creative ways of finding the answer. The unplanned cheaters often try to copy from fellow classmates, consult textbooks, use cell phones to get an answer or even access online information if invigilation is weak or non-existent. Although unplanned cheating is a kind of accidental cheating, it is still intentional and should be treated as such. Unplanned cheating cannot possibly be relevant to secure credentialing testing. There is always some invigilation in place for standardized credentialing testing and cheating detection techniques are sufficiently robust to detect copying.

Unintentional cheating can be described as any form of unfair score gain as a result of a candidate’s prior access to some of the exam materials. This could happen in many different forms.

### **What types of “unintentional” cheating occur in credentialing testing?**

One form of unintentional cheating could happen when students share information about their exams. It is common practice for students to ask previous candidates about their experience of the examination. This exercise can become a cheating issue if the same test is administered at different times. In some Asian countries many candidates share exam information through online social media. Students who take the tests will post what they can remember from the exam to a web forum; once hundreds of test takers participated in such data sharing, the entirety of a test may be reconstructed in a short period of time. A secure standardized test may then be available for free for future test takers if the credentialing testing agency does not spot the compromise made on their test. Most future candidates who see these items on the social media do not see anything wrong with memorizing them, which inevitably improves their test scores. This memorization practice has now become part of candidates’ cultures in many countries.

A different unintentional cheating may occur when candidates buy online test preparatory materials which are actually stolen test items. The preparatory test materials are sold under the warrantee of “improving your test scores”. The candidates who buy such items may genuinely think that practicing on those test items will improve their test scores. They may gain unfair advantage without knowing what they had purchased was indeed stolen test items.

The unintentional cheating is not restricted to candidates accessing test materials. In some cases teachers will try to access exam materials and teach their students to the test without candidate’s knowledge. In 2011 an undercover journalist in the UK disclosed that teachers paid up to £230 a day to attend seminars with chief examiners during which they were advised on exam questions and the exact wording that pupils should use to obtain higher marks. One of the chief examiners was caught on tape admitting that what he was doing was cheating. He told the teachers that he was telling them the cycle (of the compulsory question) and if the regulator found out he would probably have been told off. The investigation exposed a system in which some exam boards aggressively competed with one another to win business from schools. In doing so, they were accused of deliberately driving down their standards to encourage schools to sign up to them.

A fourth form of unintentional cheating is when schools are involved in the collusion of results. The individual candidates who would unfairly advantage from this act may not even know that their results were manipulated. The Atlanta scandal in 2009 is a prime example of such unintentional collusion.

### **What is the impact of this behavior on the validity of test scores?**

It is clear that scores obtained from unintentional cheating may cast doubt on the validity of the test scores and cannot be interpreted as a fair reflection of the candidates’ abilities. It goes without saying that if unintentional cheating is widespread, even honestly obtained test results may lose credibility and certificates become devalued.

### **What methods or approaches are available to detect this type of behavior?**

Geranpayeh (2014) argues that cheating is an inevitable consequence and a by-product of high-stakes testing. This is especially true of credentialing testing. Unintentional cheating will result in unusual score gains in certain schools, districts or test centers which can be detected by standard cheating detection techniques. We have already discussed that unintentional cheating is



mostly promoted through online social media. Its detection methodology should also involve monitoring such media. Many testing agencies have taken the view that it is not a matter of if the content is compromised online; it is how widespread the content sharing is. As a result they regularly monitor online social media for references to their test materials. Once a threat to their test content is identified, they decide what action to take which could be either to take legal action to close the sites or to investigate score gains of those involved in the social media forum. Sometimes they act on both. There is even some commercial software, such as CAVEON Web Patrol, that can do the job for those testing agencies that cannot afford to carry out this work by themselves.

In recent years special detection techniques have been developed to detect collusion by schools. For example, van der Linden and Jeon (2012) have developed a new technique to address answer erasure cheating that happens at school level.

### **What can test programs/owners do to prevent this type of behavior?**

The best approach to combat unintentional cheating is prevention in the first place. Many candidates or stake holders who are involved in cheating may not know that their actions are unacceptable. The Association of Test Publishers (ATP) has recommended a messaging campaign to ensure that all the stakeholders involved in the testing process are aware of the consequences of cheating. Their campaign is focused on 3 areas: test development and administration (before the exam), general communications and marketing (during the exam process), and messaging related to enforcement (after the exam). They emphasize communication of security policies to stakeholders as the best measure to prevent cheating. The ATP concludes that security messaging is a “best practice” that is critical to a program’s exam security efforts. The average stakeholder must be exposed to security messages several times before the information “sticks,” so a layered approach is recommended. Security messaging is most effective when it is woven through every aspect of the program and repeated at multiple points in the exam cycle.

### **Final Remarks**

It has already been argued that cheating is an inevitable consequence and a by-product of high-stakes testing; as the stakes of a test increase, so does the level of cheating.

One of the key assumptions in all the discussions we have had is to do with the allocation of the blame and punishing the test taker. The question is whether it is fair to punish the unintentional cheater in the same way as we do with the intentional fraudulent test taker. How do we know who willingly cheats and who doesn’t? What we know is that there is far more cheating going on in our exams than we would like to admit or are able to detect. One possible avenue for future investigation in this area is to look at cheating from a different perspective. All the statistical techniques developed for cheating detection try to identify aberrant behavior and demonstrate how improbable they may have been. If one takes the view that cheating is an inevitable consequence of high-stakes testing, why not treat cheating as a misfit, not dissimilar to rater misfit. Our current statistical procedures can identify cheating in the form of aberrant scoring patterns. Once these aberrant scoring patterns are identified, we could treat them as misfit and adjust the scores accordingly as we might do with rater misfit. If the misfit is beyond certain statistical tolerance and we have evidence that the candidates were victim of unintentional cheating, we could ask them to re-sit the test.

### **References**

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, DC: American Educational Research Association.
- Cizek, G. J. (1999). Cheating on tests: How to do it, detect it, and prevent it. Mahwah, NJ: Erlbaum.
- Cizek, G. J. (2001, April). An overview of issues concerning cheating on large-scale tests. Paper presented at the annual meeting of the National Council on Measurement in Education, Seattle, WA.
- Geranpayeh, A. 2014. Detecting plagiarism and cheating. In Kunnan, A. J. (Ed.), The companion to language assessment (pp. 980-993). New York, NY: John Wiley & Sons, Inc.
- Grijalva, T., Kerkvliet, J., & Nowell, C. (2006). Academic honesty and online courses, *College Student Journal*, 40,180-186.
- van der Linden, W. J., & Jeon, M. (2012). Modeling answer changes on test items. *Journal of Educational and Behavioral Statistics*, 37(1), 180–99.
- Wollack, J. A. & Fremer J. J. (Eds.) (2013). Handbook of test security. New York, NY: Routledge.

## Information Technology (IT) Certification

*Russell Smith, Alpine Testing Solutions*

### What types of "accidental" or "unintentional" cheating occur in Information Technology (IT) certification testing?

Many IT certification exams are on-demand and administered globally. Cheating is pervasive and it can be difficult to discern unintended cheating from malicious cheating. Cultural perspectives on what constitutes cheating vary greatly. I imagine that having prior knowledge of exam content is sometimes accidental. It may be intentional on the part of the examinee but not defined as cheating by them or it may be that training organizations are providing live test items without the candidate's knowledge.



### What are the impact(s) of this behavior on the validity of test scores?

Cheating is an absolute threat to validity. The impact of prior examination content knowledge, be it intentional or unintentional, can be dreadful. I frequently see results that appear to show two distinct dimensions being measured. One dimension is the knowledge, skills, and abilities the test is designed to measure. The other dimension is the candidates' abilities to gain an unintended advantage. Such results can call into question the meaning and interpretation of all scores.

### What methods or approaches are available to detect this type of behavior?

There are multiple ways of detecting intentional and unintentional cheating. One Catch-22 about sharing specific methodologies in the academic community is that those involved with nefarious activities pay attention to what the testing industry is saying and doing. I think because of this many organizations are hesitant to submit many of the methodologies that are being developed for peer review or present them at conferences. It is an understandable shame.

There are, however, some methodologies that are not as susceptible to such awareness. For example, comparing candidates' scores on exposed content to unexposed content using differential person functioning (DPF) analysis has proven useful for identifying candidates and groups of candidates with prior knowledge or an unintended advantage. I, along with my colleagues, have been working on ways of using DPF and differential item functioning (DIF) analysis to filter and analyze the candidates and items in ways that are meaningful and interpretable, depending on the purpose of the analyses.

Regardless of the methodology, the results should be presented in a relatively straightforward manner and fairly easily explained to various stakeholders and decision makers.

### What can test programs/owners do to prevent this type of behavior?

The single most effective thing I have seen to prevent this type of behavior is to continually refresh examination content. The rate and extent of needed refreshing depends on the extent and timing of the exposure. Windowed testing, rather than having continuously available tests, can prevent this type of behavior and is relatively cost effective. Educating candidates regarding what is considered "cheating", and what are legitimate and non-legitimate study resources, can help reduce this behavior. Delaying score reports until forensic analyses are run can deter this behavior. Finally, our primary mindset as psychometricians and test developers should NOT be to try to "catch cheaters," but rather to try to make a strong validity argument for the decisions that are being made based on the results of these exams. With that as a guiding principle, we can make good decisions regarding the development and maintenance of IT certification testing programs.

## Large-Scale Educational Assessment

*Paul M. Stemmer, Steven G. Viger, and James A. Griffiths, Michigan Department of Education*

Cheating during standardized testing has garnered a great deal of national attention. As the stakes for these assessments continue to rise, so does the incentive for schools to show significant progress, regardless of how they get there. This could increase the chance of dishonesty or cheating. In Michigan, we have seen a very low incidence of such reports and multiple methods have continued to reveal very few cases. Unfortunately, the media tend to overlook such data and focus on the few cases of potential test fraud.

Many cases the media and others outside of the assessment world tend to identify as "cheating" turn out to be accidental misadministration. Often, the problem can be traced to the proctor simply not following or misreading the directions, or using the wrong forms mistakenly. On the other hand, it is not always clear



*Pictured from left to right: Paul Stemmer, Jim Griffiths and Steve Viger*



if the school created a misadministration for the intentional purpose of improving their student outcomes (i.e., cheating). For example, there have been multiple instances where the proctor returned completed answer documents back to students, asking them to fill in all the lines on writing constructed response questions. The school would argue that they were not directly influencing the student results, because they were “not reviewing what the students wrote,” only that they filled in all the lines. The instructions in the administration manual in this case explicitly state that when students hand in the answer document and test booklet, their assessment is considered to be finished or complete. Furthermore, the materials are not to be handed back and certainly do not include any directions to the students to “fill out every line.” Is this a situation where teachers are simply encouraging students to do their best, complete work, perhaps as they do in classroom assignments/tasks? Certainly it could be - but unfortunately the behavior is strictly prohibited according to administration practices. The proctors do not see this as cheating and would likely argue that if they were cheating, they did so unintentionally, or ‘by accident’.

There are numerous other variations on the theme of accidental or unintentional misadministration. Here are a few more examples:

1. Giving students more directions than what the proctor is required to read.
2. Asking students to complete the test questions when the instructions say if they are finished to stop testing, or lets them return to online assessment or paper and pencil sections when not allowed to do so (although many online systems preclude this).
3. Putting materials about the room that would aid the student in their answers. (This can be intentional, but more frequently it is a failure to take down materials that could be aiding students in their answer.)
4. Lax monitoring of students by the exam proctors thereby allowing students to work “collaboratively” on answering their test questions.

Cheating can be defined as the intent to alter student answers to artificially improve their scores beyond what are their students’ true skills and abilities. An NCME position statement is available in another comprehensive document from the Council of Chief State School Officers - Technical Issues in Large Scale Assessment (TILSA) committee (TILSA Test Security Guidebook).

The NCME position statement defines cheating as an intentional act, but does not reference accidental or unintended cheating. Regardless of which definition one chooses to adhere to, it is clear that deviations from standard practice, such as coaching or encouraging answer completion, constitute behavior beyond what is allowed by the directions. Clearly this becomes a threat to validity. It is commonly accepted that validity is not a property of the assessments or tests - those are just the means. What is valid or not, are the inferences one makes when the instrument is used as intended. It follows that any significant deviation from standardized administration (an irregularity), would also lead to a use of the instrument beyond the scope of the intended purposes. Therefore, the inferences one can make from an instrument used under non-standard conditions are compromised and not the same as intended. Cheating, or any type of assessment practice deviating from the original plan of action, should first and foremost be approached as a validity issue.

At what point does an irregularity in administration rise to the level of invalid scores? How much evidence is required to invalidate any or all of the student scores in a given administration? This is a policy issue more than it is a measurement one. Unfortunately, in most cases (unless the school happened to be monitored) there is no external witness outside of school staff to verify what actually happened. For Large-Scale Assessment, we heavily rely on a trust between school staff as proctors to give a standard administration. When we have enough evidence that a significant irregularity or deviation from accepted practices has happened, we no longer have trust in those scores.

### **Prevention is the best defense!**

Preventing unintended or accidental misadministration is resolved by proper training and making test administration directions as clear as possible. Also, actions by the state to monitor and ensure compliance to standardized administration policies serve to prevent misadministration. Although the incidents are rather infrequent, the state recognizes that mistakes are made. There is a process in place to quickly work with schools to remediate the problem, determine the extent of the misadministration, and wherever possible, salvage the assessment results.

One form of misadministration that is becoming more common occurs when teachers expose live items to the students. A remedy, and good assessment practice, is to have a large database of test items. This is helpful in the case of computer-based administration because a large pool of items tends to inhibit the ability of students and proctors to memorize and share items (intentionally or unintentionally). In the case of paper and pencil tests, a large pool of items would allow for multiple forms of the test including emergency forms.

Adequate and comprehensive training is going to be critical. Training programs for proctors should always include examples of common irregularities and best practice. Those examples should be reinforced by stressing the critical outcome, valid vs. invalid results. Of course, prevention of irregularities should be the focus and providing some structure and focus to remediation after the fact is also important to minimize incidents and to strengthen the trust between schools and the state agency.

### **Self-reporting**

In Michigan, schools are encouraged to make self-reports of assessment irregularities as soon as they are detected. Almost always, a self-report will forgo the need for an investigation, saving the school, district, and state time and money, as well as lessening the embarrassment in the face of parents and community.

As a further incentive to prompt self-reporting of assessment irregularities, schools are well informed that doing so before the end of the test window greatly increases the chance that student scores can still be obtained through administration of an emergency test. If the school self-reports and an emergency test is prescribed, the tests are provided free of charge.

On the other hand, if a school assessment irregularity is reported by someone external to school administration during the test window and an emergency test is prescribed, the school must pay the cost of \$50 per student per content area for each emergency test scored. This policy has resulted in a significant increase in self-reports and the number of student scores that are “saved” as a result of the emergency administration.

### **Detection**

Quality assurance monitoring is practiced across the state. While the state cannot monitor all programs, the monitoring is done without advance notice and sampled in such a way that any school administration should feel they may be subject to state inspection, at any time. Additionally, schools identified in the past as having some type of irregularity, are subject to targeted monitoring.

Analyzing psychometric data for anomalous data patterns is one way to raise the question of possible irregularities. Current statistical methods are not considered sufficient to establish whether a true irregularity has occurred. If the statistical analysis suggests there is an anomaly, we first request self-investigations by the school. If the state feels it necessary, we also initiate independent investigations to attempt to accumulate further corroborating evidence. Forensic data analysis is, as of now, best used in a confirmatory framework; perhaps consider it to be one piece of the puzzle but certainly not the only or most important piece.

Although forensic data analysis is becoming more sophisticated, we are just beginning to explore ways to look at anomalous student behavior patterns on paper and pencil and computer-based assessment. We will continue to explore and develop data analysis methods for identifying anomalous patterns that are consistent with misadministration. Furthermore, the state has begun working on a database system to track historical data patterns on problematic areas for flagging results. It will take some time to develop a more sophisticated approach to historical analysis because in order to know what is abnormal we need to know first what is normal. With the variance of accidental or unintended “cheating” or irregularities being quite considerable, such a determination of normal behavior could take some time. Any results should always be subject to the proper “smell tests” to ensure that we are not merely chasing shadows best explained by statistical rationales and not supported by any practical evidence or reasoning. As more assessments are delivered online, new procedures are being developed. For example we will be able to closely monitor test sessions and data logs as well as timestamps and student behavior (e.g., telemetry and answer changing) during the assessment. Doing so in real time, with faster and more robust data storage and transfer solutions, will make analytics even more effective and perhaps be better utilized in the proactive rather than the reactive sense.

### **Investigations**

Much like flagging anomalous statistical results, if we have an anonymous tip or complaint, or we have a strong suspicion that a school has an irregularity the school may be requested to file a self-report investigation, or the state may immediately launch an investigation. There is a toll free hotline to reach the Division of Accountability Services, which provides access to a tip line as well as a website reporting location.

Anything reported is quickly acted upon and the proper internal staff is notified: Integrity Officer, Reporting and Administration Manager, Chief Psychometrician, Assessment Director. If the situation is deemed to potentially be a serious threat to validity, independent trained investigators will go to the school to conduct a fact-finding. The goal of the investigation is to find all the facts regarding how the administration was conducted. This is accomplished by interviewing the teachers (all personnel who served as proctors), the test takers (students), building administrators and other staff involved in assessment. All memos and training information housed at the school is also reviewed by the investigators. The overall goal of these practices is to gather evidence for or against the complaint and gather evidence of whether standardized practices were properly followed.

## **Remediation**

Remediation in Michigan consists of requiring emergency tests if time allows. Invalidation of the scores is sometimes necessary. The state may also require the school staff to undergo more training, file a training plan, and other methods to prevent the problem from reoccurring. The discipline of staff by the local district that caused the irregularity can range from reprimands, additional training requirements, or dismissal.

## **Finding Policies that Work**

Unfortunately there will be examples. As allegations of potential test fraud become more frequent and public, we hope it serves as motivation for the schools to work more closely with the state to prevent misadministration and irregularities and to remediate them during the assessment windows.

Laws in some states require those found to have engaged in cheating or other behavior deemed inappropriate enough to invalidate scores to face suspension or full removal of teacher credentials. Proving the question of intentionality requires a more extensive investigation. Michigan does not have a law that requires the state to take actions against a school or teacher for cheating. A determination of intentional improper staff behavior is left to the local to school to be investigated and dealt with. In many cases the local district is closer to the problem and can deal with personnel issues in a more appropriate way. MDE cannot say with 100% certainty that administrators and/or teachers have been disciplined as a result of our reports but we have good evidence this has happened in some cases. It is good practice to carefully consider different policy models and continuously improve them.

---

## **Update from the Standards and Test Use Committee**

*Rosemary Reshetar (Chair), The College Board*

A few years ago, before joining the STUC, I wasn't aware the committee existed much less what they did. Here we are, four years later. I'd like to use my newsletter space to give an update on the committee's activities and future plans. First, the committee consists of seven members typically including three from academia, three from industry, and one graduate student representative.



When I joined the committee the Test Standards were undergoing revision, and there was work to be done. Our role was to facilitate a comprehensive review by NCME members of each chapter and compile minor and major suggested revisions to present to the NCME Board who would then present recommendations on behalf of NCME to the Joint Committee. The Joint Committee then reviewed and considered the responses obtained from each of the three organizations (AERA, APA, NCME). Fast forward to today and we expect the latest version to be available for purchase at this year's annual meeting.

Moving forward this committee plans to undertake some less visible, though important, maintenance and update projects to serve the measurement community. For example, as the committee is responsible for monitoring professional issues concerning the development, validation, and proper use of assessments, we plan to focus on review and dissemination of information regarding other standards and guidelines for test use. While some of these publications are available through the NCME website, a review with possible additions is in order. Beyond simply collecting and disseminating relevant documents, this committee also maintains the policy on NCME's sponsorship, approval, and endorsement of products. Part of our work will include incorporating these aspects into the review process, making any procedural updates if necessary.

In closing, please join me in welcoming Phillip Ackerman as the incoming committee chair.

---

## NCME ANNUAL MEETING UPDATE

The NCME annual meeting is just weeks away! Here is some important information as you make your final preparations.

### Invited Speaker Sessions

Make sure to mark your calendars and attend the invited speaker sessions:

Friday April 4 11:50 -12:40

Ryan Baker – technology and computer science

Angela Duckworth – psychology

Saturday April 5 11:50 - 12:40

Donald Hedeker – statistics

Valerie Shute – education measurement

Sunday April 6 11:50 - 12:40

Erick Hanushek - economics

Jim Popham – education measurement

### Important Links

Annual meeting website: <http://ncme.org/annual-meeting/next-meeting/>

Meeting Registration (through AERA):

<http://www.aera.net/EventsMeetings/AnnualMeeting/2014Registraton/tabid/15262/Default.aspx>

Hotel and Travel Information (through AERA):

<http://www.aera.net/EventsMeetings/AnnualMeeting/2014HousingandTravel/tabid/15278/Default.aspx>

## NCME FITNESS WALK/RUN CORNER

*Brian French & Jill van den Heuvel for the NCME Fitness Walk/Run*

The details for the NCME Run/Walk are almost finalized. We have neat tech shirts again this year with the design selected by prior participant survey respondents. Please pick up your shirt and information at the NCME hotel at the main NCME information table. The event will be a short walk from the hotel, so no need for buses this year!

Just meet in the lobby of the NCME hotel by 5:45am Sunday morning. There is still time to register on-line or on-site through the AERA main registration process. Be on the lookout for an email soon with additional details if you have already registered.



## PHILADELPHIA TRAVEL

*In advance of our upcoming visit to Philadelphia, a couple of NCME members offered to share their local favorites for things to do, see, and eat!*

### Philly Phun

Bill Herrera, edCount, LLC



#### Places to Visit

- The Pennsylvania Convention center is very close to the Reading Terminal: Good for lunch and shopping around for Amish wares (<http://www.readingterminalmarket.org>). Also in there is the Flying Monkey Bakery (<http://www.flyingmonkeyphilly.com>).
- Brew tours: Yards Brewing Company (<http://www.yardsbrewing.com>) and Philadelphia Brew Company (<http://philadelphiabrewing.com>)
- For a carousel ride and Mini-golf, go to Franklin Square (<http://historicphiladelphia.org/franklin-square/what-to-see/>).

#### Museums

- The Barnes (<http://www.barnesfoundation.org>)
- Pennsylvania Academy of the Fine Arts (<http://www.pafa.org>)
- Mutter Museum (<http://www.visitphilly.com/museums-attractions/philadelphia/mutter-museum/>)
- Edgar Allan Poe (<http://www.visitphilly.com/history/philadelphia/edgar-allan-poe-national-historical-site/>)
- Academy of Natural Sciences (<http://www.ansp.org>)

#### Restaurants and Bars

- You can't go wrong with any from Steven Starr. He has all price points and all kinds of foods (<http://www.starr-restaurant.com>).
- In Chinatown, some favorite places are Vietnam Palace (<http://www.vietnampalace.net>) and Penang (<http://www.yelp.com/biz/penang-philadelphia>).
- For a speakeasy, Hop Sing Laundromat (<http://hopsinglaundromat.com>) or Franklin Morgan and Investment Co. (<http://www.thefranklinbar.com>).
- For ice cream, Franklin Fountain (<http://www.franklinfountain.com>).

#### Coffee

- Olde City Coffee in the Reading Terminal (<http://www.oldcitycoffee.com>)
- Town Hall (serving Reanimator, <http://townhallcoffee.com/location/center-city-cafe/>)
- HubBub coffee (<http://hubbubcoffee.com/wordpress/>)

#### Concert Venues

- Union Transfer (<http://www.utphilly.com>)
- Johnny Brenda (at my house, <http://www.johnnybrendas.com>)
- Boot and Saddle (<http://www.bootandsaddlephilly.com>)
- The Trocadero (<http://www.thetroc.com>)

This is always a good site to be in the know that weekend, too: <http://www.uwishunu.com>.

## Some Philly Favorites

Andrew Wiley, *Alpine Testing Solutions*



### Places to Visit

Philly has some great museums for people to consider. Here are a couple to consider, aside from the Art Museum, which has the “Rocky” steps.

- Barnes Museum (<http://www.barnesfoundation.org/>)  
The Barnes Museum has an amazing collection and only recently opened up in downtown Philly. For years, the collection was housed in a museum in the distant suburbs and only open one day a week. It is definitely worth a visit.
- Rodin Museum (<http://www.rodinmuseum.org/>)  
If you like sculpture, this is a great collection. They claim it is the largest collection of Rodin’s work outside of Paris.

A couple slightly different options:

- The Mutter Museum (<http://www.collegeofphysicians.org/mutter-museum/>)  
A fairly crazy collection of old medical devices, anatomical freak shows, and just strange stuff you won’t find anywhere else. Definitely worth a visit if you are looking for something out of the ordinary.
- Eastern State Penitentiary (<http://www.easternstate.org/>)  
Who wouldn’t want to tour an old prison? Former housing for Al Capone and many others. Pretty wild architecture and rumored to be haunted.

### Get Outside

Philadelphia is quite proud of its series of murals and artwork throughout. Here are a couple of great ones:

- Museum without walls (<http://museumwithoutwallsaudio.org/#>)
- Mural arts program (<http://www.muralarts.org/>)
- Go for a bike ride along the Schuylkill River. You can get from downtown Philly to Kelly Drive in about a mile, and it is a great place to relax and get some exercise.

### Food

For food, everyone talks about cheesesteaks in Philly. But I would recommend you find a place that serves a good pork loin sandwich. Two of the famous ones:

- DiNic’s in Reading Terminal (always crowded, so go early)
- Tony Luke’s in South Philly



## AERA – DIVISION D LUNCHEON

*Division D Thanks  
The Sponsors of  
the 2014 Luncheon and  
Business Meeting*



**BUROS**  
CENTER FOR TESTING



*Listening. Learning. Leading.®*



**PEARSON**

### *AERA Division D*

*Vice President*  
**Wayne Camara**  
ACT

*Secretary / Newsletter Editor*  
**Kristen Huff**  
Regents Research Fund

### *2014 Luncheon and Business Meeting Committee*

**Javarro Russell, Chair**  
National Board of Medical Examiners

**Rochelle Michel**  
Educational Testing Service

**Amy Hendrickson**  
College Board

**Richard Feinberg**  
National Board of Medical Examiners

**Jing Mao**  
Educational Testing Service

**Amanda Soto**  
National Board of Medical Examiners

### *2014 Division D*

*Measurement  
&  
Research Methodology*

*Luncheon and Business Meeting*

**Friday, April 4, 2014**  
**12:25 p.m. – 1:55 p.m.**

Marriott, Third Level - Liberty ABC  
Philadelphia, PA



**AERA** AMERICAN  
EDUCATIONAL  
RESEARCH  
ASSOCIATION

### *Keynote Speaker*



**Dr. Alina von Davier,**  
*Director of Research*  
ETS

### *Agenda*

- ✚ Welcome & Introductions
- ✚ Thank You to Sponsors
- ✚ Awards and Certificates of Appreciation
- ✚ Review of 2014 Budget
- ✚ Luncheon Speech:  
“Test Score Equating and Measurement Models” by **Dr. Alina von Davier**
- ✚ Pictures with Awardees and Committee Chairs

### *Award Winners*

- ✚ Robert L. Linn Distinguished Address Award:  
**Dr. Eva Baker**, *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*
- ✚ Significant Contribution to Educational Measurement and Research Methodology:  
**Dr. Howard Wainer**, *National Board of Medical Examiners*
- ✚ Early Career Award in Measurement and Research Methodology:  
**Dr. Jessica Nina Lester**, *Indiana University*
- ✚ AERA Division D Outstanding Quantitative Dissertation Award:  
**Dr. James Pustejovsky**, *University of Texas at Austin*

---

## INDUSTRY EVENTS

### **CASMA Equating Workshop, June 9 – 13, 2014 --- Iowa City, IA**

A more intensive and extensive workshop than a one-day session on equating will be offered by CASMA, June 9-13, 2014. Five similar workshops were held previously. They were well received and well attended. No workshop can replace a full-length course, but this particular workshop should provide participants with a good working knowledge of basic equating designs, statistical procedures, and applications.



Visit the workshop web page for details: <http://www.education.uiowa.edu/centers/casma/workshops>

Should you have any procedural, housing, or registration questions, contact Jennifer Jones at 319-335-5954. Workshop content questions can be directed to Bob Brennan, University of Iowa, at 319-335-5405.

### **Northeastern Educational Research Association (NERA)**

The 45th Annual NERA Conference will be held October 22-24, 2014 at the Marriott Hotel in Trumbull, Connecticut.

Proposals *must* be submitted electronically by June 6<sup>th</sup>, 2014.

The Call for Proposals is attached here  along with the Frequently Asked Questions 

---

## **Call for Nominations: Editor, NCME Newsletter**

Deadline: April 15, 2014

The NCME Publications Committee is soliciting nominations for Editor of the *NCME Newsletter*, to serve a three-year term beginning January 2015. The Publications Committee will screen nominations and offer a slate to the President and NCME Board of Directors, which make the appointment in early summer 2014.

The *NCME Newsletter*, published quarterly, includes announcements and brief descriptions of current activities of interest to the membership, including significant publications, upcoming meetings, and NCME Board and committee activities. The *Newsletter* provides NCME members with timely information about current events in measurement practice and research, including news from state and federal agencies, school districts, universities and colleges, credentialing agencies, educational research laboratories, and publishers.

Major responsibilities of the Editor:

- Plan content for each issue, including regular columns and feature articles.
- Solicit, review, and edit content for the Newsletter.
- Write material for each issue (e.g., regular column); correspond with authors as necessary.
- Communicate with the President and Board to identify and summarize newsworthy association activities.
- Serve a three-year appointment from January 2015 through December 2017; report to the NCME Board through the Publications Committee.
- Appoint an Editorial Board to assist with all activities.

A modest stipend is available to help support the position in accordance with NCME policy. If interested in this position, or if you would like to nominate a colleague, please contact Mark Raymond, Publications Committee Chair, by email ([mraymond@nbme.org](mailto:mraymond@nbme.org)).

**Deadline for nominations is April 15, 2014.**

---

## In Memory of Howard C. Mitzel, Ph.D. (1951 ~ 2014)

*Prepared by Alan Nicewander, Pacific Metrics*

Monterey, CA – Howard C. Mitzel, Ph.D., principal founder of Pacific Metrics Corporation in Monterey, CA, passed away on January 19, 2014. He was born in 1951, the son of Harold E. Mitzel and Lois Howard Stuart. His father, Harold, served as Executive Editor of the *Journal of Educational Research*.



Howard held a Ph.D. from the University of Chicago in psychometrics and cognitive psychology. He held three software patents and was the co-inventor of the Bookmark method for setting standards and cut scores for educational tests. Howard developed and programmed an ensemble of psychometric and statistical software that is used for every test developed at Pacific Metrics. Dr. Mitzel co-founded Pacific Metrics in 2000 to advance the transition of technology for educational assessment into the American classroom. He led the development of an automated scoring system for essays and short answer test questions, which was a prize-winner in a 2011 education technology competition. Prior to forming Pacific Metrics, he worked as a Senior Research Scientist at CTB McGraw-Hill. There, he co-developed Bookmark, the most widely used standard setting method in K-12 education in the United States.

Howard's influence in the measurement field went beyond his substantial technical and methodological contributions. He served on the technical advisory committees to several state departments of education and was widely known—and greatly respected—in the testing community of the US. Howard's integrity was well known and greatly admired, and it reflected the importance he placed both on personal ethics and on professional responsibility. Two examples illustrate the role of integrity in Howard's life: (1) as a matter of ethics, he refused to use his advisory role to state boards of education as a means of selling Pacific Metrics' products and services to the states for which he was an advisor; (2) he also refused to allow Pacific Metrics to participate in (unvalidated) teacher evaluation systems that used student test scores. In fact, Howard is the co-author of a letter to appear soon in *Education Week*, in which a critique is leveled at the evaluation of teachers using student test scores; he points out that the use of student scores, for this purpose, would constitute a massive violation of the *APA-NCME Standards for Educational Tests*—unless the appropriate validity studies are conducted prior to implementation.

Howard was also known for his publications and presentations at educational meetings on standard setting; the use of automated scoring for essays and short-answer questions; the use of automated scoring in the detection of “scoring drift” in human ratings of essays; and contributions to many other important technical and policy issues in educational assessment.

It would not be an adequate review of Howard Mitzel's life without mentioning his wonderful sense of humor, which included elements of both sophistication and naughtiness. For example, one evening after work he was sitting with a male colleague at a classy Monterey bistro. An attractive woman came by and introduced herself, then practically fell into the lap of his male friend while straightening some pictures on the wall behind him. His friend noticed from the corner of his eye, that Howard was busily making crooked the pictures on the wall beside him.

In the end, the dreadful cancer that had taken his voice ended his life, but not before Howard had fought courageously—never losing his sense of humor or the hope that he was going to win the battle. His surgeons at Stanford University Hospital often remarked that he showed great courage and a tremendous will to survive.

Dr. Mitzel is survived by his mother, Lois Stuart; his sister and brother-in-law, Claudia and Richard Kerbel; niece, Allison Kerbel; and nephew, Nathan Kerbel. Memorial contributions in his name may be made to: National Public Radio ([www.npr.org](http://www.npr.org)) and ASPCA ([www.aspc.org](http://www.aspc.org)).

**To get the NCME Newsletter four times a year  
(March, June, September, and December)**

**Go to: <http://ncme.org/publications/newsletter/>**

*Newsletter Advisory Board*

**LANIE BRADSHAW**, University of Georgia  
**JOHN DENBLEYKER**, Minnesota Department of Education  
**ELLEN FORTE**, edCount  
**JOANNA GORIN**, Arizona State University  
**SARA HENNINGS**, MetaMetrics, Inc.  
**JOAN HERMAN**, CRESST/UCLA  
**ANDREW HO**, Harvard Graduate School of Education

**THEL KOCHER**, Walden University  
**GERRY MELICAN**, The College Board  
**MELINDA MONTGOMERY**, University of Kansas (Grad Student Rep)  
**THANOS PATELIS**, Center for Assessment  
**S.E. PHILLIPS**, Consultant  
**CHRISTINA SCHNEIDER**, Center for Assessment  
**XIAN (BO) WANG**, The College Board

**SUSAN DAVIS-BECKER**, Editor, Alpine Testing Solutions

*Send articles or information for this newsletter to:*

Susan Davis-Becker  
Alpine Testing Solutions

Phone: 402-483-5898  
e-mail: [susan.davisbecker@alpinetesting.com](mailto:susan.davisbecker@alpinetesting.com)

The *NCME Newsletter* is published quarterly. The *Newsletter* is not copyrighted; readers are invited to copy any articles that have not been previously copyrighted. Credit should be given in accordance with accepted publishing standards.