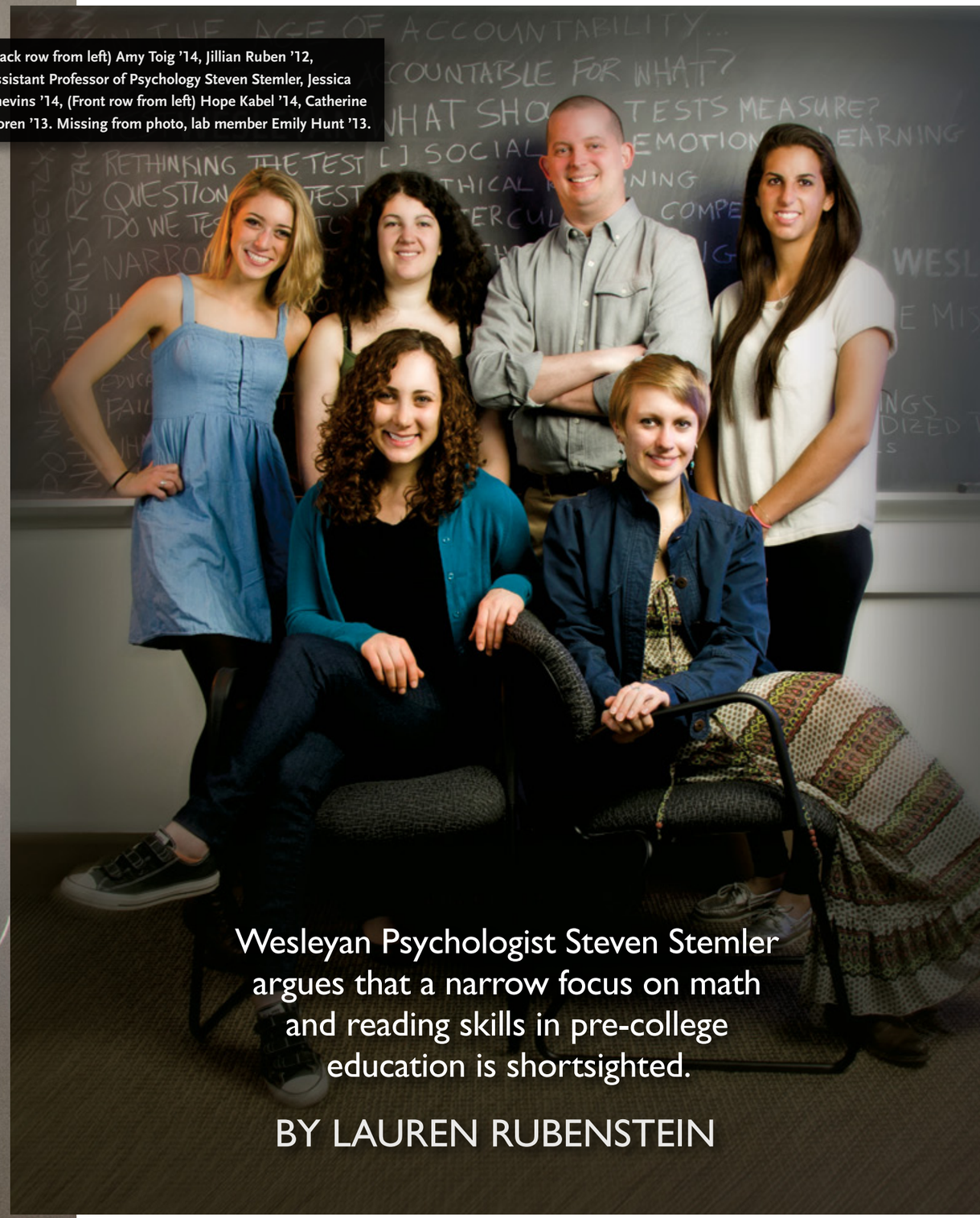
Test QUESTIONS

PHOTOGRAPHY BY JOHN VAN VLACK



(Back row from left) Amy Toig ’14, Jillian Ruben ’12, Assistant Professor of Psychology Steven Stemler, Jessica Shevins ’14, (Front row from left) Hope Kabel ’14, Catherine Doren ’13. Missing from photo, lab member Emily Hunt ’13.

Wesleyan Psychologist Steven Stemler argues that a narrow focus on math and reading skills in pre-college education is shortsighted.

BY LAUREN RUBENSTEIN

Since the enactment of the landmark No Child Left Behind Act a decade ago, "accountability" has been the buzzword in our nation's education policy. Schools are held accountable for student learning, as measured by standardized tests in reading and math, administered annually to every student in grades three through eight and once in high school. This testing has been the subject of fierce national debate since the law's passage, with many wondering whether the negative impacts outweigh the positive.

But Assistant Professor of Psychology Steven E. Stemler poses a novel question: Are we even testing for the right things?

Any examination of school accountability must consider the role schools fill in our society. In a newly published book, *The School Mission Statement: Values, Goals, & Identities in American Education,* Stemler and co-author Damian J. Bebell of Boston College contend that a narrow focus on students' math and reading abilities is shortsighted. After all, the purpose of school has never been simply to improve students' cognitive abilities—though that is, of course, a major goal. Rather, schools have long sought to develop a wide range of competencies in students, ranging from social and emotional learning and ethical reasoning to intercultural competence and citizenship. Shouldn't schools be held accountable for student growth in these areas as well, Stemler and Bebell ask.

Or, as they put it in the book: "The current accountability system is like evaluating the effectiveness of basketball players exclusively by how high they can jump: Although jumping is an important, if not critical, component to a basketball player's success, it fails to represent all of the other important components that make for a great teammate, strategist, etc."

From state constitutions and court decisions to the musings of philosophers and historians, there has been much written on the purpose of school. Stemler and Bebell discovered that these various sources all tend to converge on a finite set of key goals for schools to pursue: civic, emotional, social, cognitive, and vocational development of students. None of the sources they reviewed gave precedence to any one purpose.

"If you look at legal or constitutional documents, there's no inherent preference given to the cognitive domain," says Stemler. "There's nothing privileged about math and language arts."

In the book, Stemler and Bebell focus on school mission statements for evidence of priorities. Schools, like businesses and civic and nonprofit organizations, use mission statements to summarize their core goals, through concise and simple statements that communicate broad themes. Past research has shown that the most effective schools tend to commit to a shared mission. Mission statements are typically created through a collaborative process involving students, teachers, administrators, parents, and community members, and thus reflect the values of local stakeholders. And school principals tend to point to the mission statement as an important tool for shaping school vision and practice.

Stemler and Bebell have developed a coding rubric for school mission statements, classifying them according to broad themes that touch on cognitive, social, and emotional development; vocational preparation; integration into the local community; integration into the global community; fostering a challenging environment, and more.

They used the rubric to analyze 111 mission statements from a number of different types of schools, ranging from public schools, vocational and technical schools, and magnet schools to parochial, Waldorf, Montessori, and other specialty schools. They concluded that most American schools share certain priorities, namely cognitive and emotional development of their students. Other values are more distinctive to particular types of schools. For example, the most frequently observed theme in public high school mission statements is civic development, but not a single Montessori school in their sample mentions it. Montessori schools prioritize emotional development and the provision of a safe and nurturing environment. And while KIPP (Knowledge Is Power Program) charter schools, which serve primarily urban, African-American populations of children, emphasize college and career preparedness, the Waldorf Schools, which typically serve a more suburban, white, affluent community, focus on finding joy in learning and purpose in life.

Stemler says he doesn't intend to pass judgment on the different values expressed in the schools' mission statements. Rather, he thinks mission statements can serve as an important tool for parents in choosing a school for their children that is consistent with their own priorities.

"I'm quite impressed with the diversity of the U.S. schools and the fact that there are all these niches out there, serving different populations. Not everybody has the same point of view on how they want schools to serve their kids," Stemler says. "I support schools determining their own priorities. They just need to be able to demonstrate that they're fulfilling the goals they articulate."

At present, though, the vast majority of school systems do not perform broad assessments that reflect their goals. While specific educational policy prescriptions are not his main focus, Stemler says that the use of well-crafted tests is essential.

But how can something as seemingly abstract as a student's level of creativity or ethical reasoning be measured?

Developing such tests is precisely how Stemler spends his days in Wesleyan's Psychometric Laboratory on Human Intelligence & Social Behavior. His lab has developed assessments to measure creativity, citizenship, ethical reasoning, intercultural competence, and social intelligence. They are geared toward a college-aged audience but could be adapted for grade school use. According to Stemler, these tests try to get at the ABCs—attitudes, behaviors, and cognitions.

For example, in a test of citizenship, participants are asked about behaviors, such as whether they belong to any political organizations or have helped out during an election campaign in the last year. They are tested on knowledge with questions such as, "How many seats does Connecticut get in the U.S. House of Representatives?" The test also provides hypothetical situations to tease out participants' commitment to certain principles associated with citizenship. For example:

It's Election Day, and there's a tight race between two candidates. Your friend, who supports a different candidate than you, wants to vote and needs a ride to the poll. Do you give him a ride in your car?

Stemler describes another test his lab developed, which measures creativity across four dimensions: problem identification, problem solving, imagination, and frame shifting, or thinking about questions in a different way.

In one question, participants are provided with three words that have no meaning in the English language. They are asked to make up definitions for the words and to communicate their meaning to someone who has similar interests and expertise, using a medium of their choice (e.g., drawing a picture or cartoon, writing an essay, poem, computer program, or mathematical proof). After that, participants are told to imagine that they want to make the new words catch on as part of everyone's everyday language. Their task is then to sketch out how they would use technology (e.g., Twitter, Facebook, YouTube) to do so. Finally, participants are told that they have been asked to present these new words to the board of Merriam-Webster to determine whether they can be officially included in the dictionary. The board accepts only words that can be understood in context without needing a formal definition. These different scenarios are testing participants' ability to frame shift.

Responses are graded based on their effectiveness in communicating the meaning of the words (as judged by independent raters), the complexity of the methods of communication, their originality (how statistically rare the ideas are compared with the responses of other test-takers), and their flexibility across scenarios (how categorically different the methods of communication used are).

Stemler and the students in his lab are now evaluating the validity and reliability of the tests they developed by administering them to a small group of Wesleyan students and analyzing the data. They are considering whether the results are consistent over time and how they relate to other measures. For example, how does someone's score on the creativity test relate to self-ratings of creativity, grades in school, life satisfaction, and creative achievements? Once this evaluation phase is complete, Stemler aims to administer the tests to a much larger group of students at Wesleyan and to students at other area universities, before adapting them for a high school population. He has been in discussions with several K–12 educators about how to apply his ideas in a grade school setting.

One such educator is Cynthia Crimmin, a longtime teacher and school administrator who serves as principal of an elementary school in Watertown, Mass., and is also president of MASCD, the Massachusetts Association for Supervision and Curriculum Development. She is working with colleagues to study creativity. Crimmin says that pre-collegiate educators are increasingly aware of the relationship between good instruction, clarity of learning objectives, and evaluation. Assessment, she says, drives learning.

"As we attempted to create rubrics for use in our schools, we discovered few examples of measures of what are being termed '21st-century skills,' such as innovation, collaboration, perseverance, and global awareness," she says. "Having a researcher of Dr. Stemler's stature take on the challenge of designing indicators to measure these attributes is very exciting to us. We agree with him that the time is ripe for this kind of initiative."

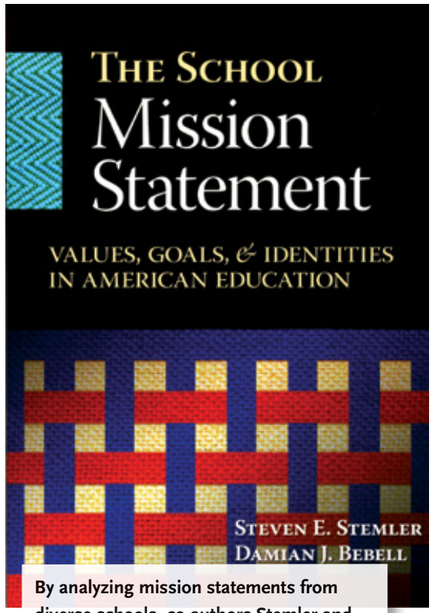
Stemler acknowledges that these tests might not yet be perfect. However, he says this is equally true of tests measuring students' cognitive abilities.

"I'd rather see imperfect tests of these broader skills put in place than have these important competencies be ignored altogether," he says.

Robert Connor, a colleague of Stemler's who has a degree in psychometrics and is head of an urban private school in Trenton, N.J., says, "Teachers are clamoring for innovation within the area of assessments. Generally speaking, what they've been given are tools that allow them to compare their kids with kids in other schools or counties. I think they're interested in tools that allow them to understand on a more granular level how their kids are learning from Monday through Friday—how they're moving forward and grasping complex aspects of the curriculum."

Dan McNeish '11 transferred to Wesleyan his junior year from a state school in Michigan because he wanted to study psychometrics, and Wesleyan was one of only a few schools to offer that opportunity to

"The current accountability system is like evaluating the effectiveness of basketball players exclusively by how high they can jump."



By analyzing mission statements from diverse schools, co-authors Stemler and Bebell concluded the institutions serve multiple purposes beyond the cognitive component for students.

undergraduates. He immediately enrolled in Stemler's Psychological Measurement course and approached Stemler about working in his lab. There, McNeish assisted Stemler with developing assessments for social intelligence, ethical reasoning, and adaptability for military officers.

"Professor Stemler works with concepts that aren't easy to define, which makes his work so interesting," McNeish says. "People throw around terms like 'creativity' and 'ethical reasoning' all the time, without really considering what they mean. But to develop a test for creativity, you first have to figure out how to define it and, then, how to measure it."

After graduating from Wesleyan, McNeish entered a graduate program in Measurement, Statistics and Evaluation at the University of Maryland, one of the top programs in the country. He said most psychometricians' work revolves around concrete concepts—like measuring math and reading ability—and that Stemler is rare in his efforts to go after these "difficult, messy concepts."

Maureen Scudder, an experienced educator who is the director of communications at the Cobb School, a Montessori School in Connecticut, agrees that by and large teachers are concerned with developing "the whole child" and have an "open-minded and healthy" attitude toward the notion of assessing skills beyond reading, writing, and math ability. She says it tends to be parents—and, as a result, administrators—who embrace traditional tests that are more limited in scope.

"If the parents were crying for a kinder, gentler, community-minded child who knew himself, really knew himself—what he felt, what he believed, what he valued, what he excelled in, what he struggled at, and on and on—then this would become a priority. But they're not. They're simply not. Today's parents are the first and greatest obstacle to creating broader assessments," she says.

Scudder adds, "I think all constituents should focus on the school's mission statement and ask the tough question: 'Is this school doing what it promises?' If everyone holds fast to the mission, then the school and its children—the most important beneficiaries—can fly."

However, she says, "I'm not sure about testing children on the tenets of the mission statement. I think they should learn it, live it, wrestle with it, question it, etc., but I don't think they need to take a test on it."

Stemler believes that broadening assessments to reflect the full range of schools' priorities would have beneficial effects beyond holding schools accountable for their stated goals. Equipped with a more complete picture of their students' strengths and weaknesses, teachers would be able to improve instruction. They could also group students with different strengths to work together on projects, so the students could learn from one another.

In addition, teachers tend to have their own ideas, based on daily classroom interactions, about which students excel in certain areas and which have deficits. Assessments can serve to either challenge or confirm teachers' beliefs. "The point of testing is to overcome biases. You may identify someone who is much more creative than you thought," says Stemler.

Moreover, he says, "When you design tests that measure these broader processes, what we find is that you reduce ethnic differences in achievement. The black-white gap that you tend to see on tests of analytical thinking ability is greatly reduced on tests of creative thinking. With tests of practical thinking, the Latino students tend to score pretty close to white students."

Overall, broader assessments help to identify strengths in students who may not shine academically, allowing teachers to "identify a broader base of talent," Stemler says.

And though our recent national education dialogue has been sharply focused on accountability in reading and math, Stemler points out that this wasn't always the case. American schools have shifted their emphasis over the decades, reflecting national trends. During the immigration boom at the turn of the 20th century, schools were viewed as places to integrate new immigrants into American society. The mid-20th-century Space Race brought a major push for math and science education, while the Self-Esteem Movement of the early 1980s shifted the focus to emotional learning. The current emphasis on cognitive development has been around since the birth of the Standards Movement in the 1990s.

High school principals told Stemler that their school mission statements tend to be revised frequently to reflect changing priorities. In a study not yet published, but which was presented at the New England Educational Research Organization in 2007, Stemler and a former student, Lauren Sonnabend '08, analyzed a sample of about 30 Massachusetts high school mission statements between 1996 and 2006, a decade split by the passage of No Child Left Behind. They found consistent evidence that schools narrowed their mission statements during this time to emphasize cognitive development. Schools that once vowed to provide a "joyful, caring, fulfilling and safe environment" now promised "world-class standards" and achievement in math and reading.

"These things do cycle through. I think we will come to the end of the cognitive cycle fairly soon, and something new will emerge," Stemler predicts.

Do you have an opinion about this topic? Please write us at letters@wesleyan.edu.

STEMLER LAB CREATIVITY TEST SPRING 2012

PART I (2 MINUTES) 1

Below are three words with no current meaning in the English language. Your first task is to define these words in any way you wish. Please indicate in the space next to the word what you have decided each of these words means. Once you have settled on their definitions they will remain constant for the remainder of the test.

AZONITY _____

IMPIRTENE _____

EMFOINKER _____

PART II (15 MINUTES) 2

Now, we would like you to convey the meanings that you have given to these words in a creative way to a reader who does not know how you have defined them. We encourage you to use any method of communication that caters to your strengths. For instance, if you are most comfortable communicating through a mathematical proof, you could convey the meanings of these words through a proof. Do not simply define the word, but rather try to convey its meaning through context.

You may use Microsoft Word and/or the paper and pen that we have provided. You may use the words as many times as necessary to accurately convey their meaning.

Your score will be determined based on the cleverness, clarity, and originality of your methods of communication, and your success in relating the three words to one another.

PART III (10 MINUTES) 3

Your next task is to try to make your words a part of everyday language. Think about how you might use various forms of advertising media (such as a billboard, a radio advertisement, an Internet ad, etc.) and/or social media (such as Twitter, Facebook, Web comic, memes, etc.) to convey the meanings of these words in a way that would make others use them, spreading their presence and popularity as far as social networking allows.

For this part of the test, please do not define the words in relation to one another, but instead think about how to convey the meaning of each word separately. Keep in mind that when a person you have never met comes across your newly defined word, he or she should be able to understand exactly what each word means and how it would be used after seeing your creation.

PART IV (10 MINUTES) 4

Because of your success in popularizing your words, you have been asked to present the new words to the board of Merriam-Webster to determine whether they can be officially included in the newest edition of the dictionary. However, the board will only accept your words if you can prove that they can be understood in context, without the need for a formal definition. Using Word and/or the paper and pens provided, draft a presentation in which you convey the meanings of the words and relate them to each other, without ever explicitly defining them.