Article

SOCIETY FOR THE TEACHING OF PSYCHOLOGY

Teaching of Psychology
1-11

DOI: 10.1177/0098628315603059
top.sagepub.com

SAGE

What Is a Psychological Misconception? Moving Toward an Empirical Answer

D. Alan Bensley[1] and Scott O. Lilienfeld[2]

Abstract

Studies of psychological misconceptions have often used tests with methodological shortcomings, unknown psychometric properties, and ad hoc methods for identifying misconceptions, creating problems for estimating frequencies of specific misconceptions. To address these problems, we developed a new test, the Test of Psychological Knowledge and Misconceptions, administering it to a sample of 162 graduate and undergraduate psychology students and to a second sample of 173 undergraduate psychology majors. Results revealed high consistency in item response frequencies across samples, allowing identification of specific misconceptions. We found that certainty ratings of the correctness of more frequently endorsed misconceptions were significantly higher than misconceptions endorsed less frequently while some other items endorsed much less frequently also showed higher certainty. Our findings bear implications for dispelling erroneous but confidently held claims in psychology courses.

Keywords

psychological misconceptions, knowledge of psychology, metacognitive monitoring, conceptual change, learning outcomes assessment

Many studies have found that students are highly susceptible to psychological misconceptions (e.g., Brown, 1983; Gardner & Brown, 2013; Kowalski & Taylor, 2009; Lamal, 1979; McKeachie, 1960; Standing & Huber, 2003; Taylor & Kowalski, 2012; Vaughan, 1977). For example, many students incorrectly believe that people only use 10% of their brains (Higbee & Clay, 1998). Educators are justifiably concerned about psychological misconceptions not only because they run counter to the scientific knowledge they are trying to impart, but because standard instruction is often not effective in eliminating them (Best, 1982; Gardner & Dalsing, 1986; Landau & Bavaria, 2003; McKeachie, 1960; Vaughan, 1977). Understanding misconceptions and their resistance to elimination is important to the scientific study of conceptual change (diSessa, 2006). Nevertheless, many misconceptions studies have used tests with methodological and psychometric shortcomings, raising doubts about the prevalence of misconceptions. We report on our initial efforts to remedy these problems.

We first examine definitions of psychological misconceptions and discuss challenges in the measurement of misconceptions. Then, we report results from a new test called the Test of Psychological Knowledge and Misconceptions (TOPKAM) designed to address shortcomings in previous measures and allow assessment of students' metacognitive monitoring of their knowledge of psychology (Bensley & Lilienfeld, 2010). We argue that using the TOPKAM with a conservative statistical selection procedure allows us to identify specific psychological misconceptions. Finally, we apply these results to test whether students are more confident of their misconceptions or of their correct responses.

Definitions of Psychological Misconceptions

Researchers have typically defined scientific misconceptions in general terms. For example, diSessa (2006, p. 269) defined misconceptions as "false, persistent beliefs" contradicted by established scientific evidence (see Taylor & Kowalski, 2004, for a similar definition). More recently, Taylor and Kowalski (2014, p. 259) defined a misconception as "inaccurate prior knowledge."

Defining misconceptions as inaccurate prior knowledge is consistent with the inference that when students endorse misconceptions on items, they are exhibiting faulty knowledge, lacking accurate knowledge, or both. Likewise, definitions that define misconceptions as preconceptions (Morrison & Lederman, 2003) and as alternative conceptions incongruent with disciplinary thinking (Piquette & Heikkinen, 2005) imply a deficiency in scientific knowledge.

[1] Department of Psychology, Frostburg State University, Frostburg, MD, USA
[2] Emory University, Atlanta, GA, USA

Corresponding Author:
D. Alan Bensley, Department of Psychology, Frostburg State University, Frostburg, MD 21532, USA.
Email: abensley@frostburg.edu

The focus on prior knowledge in defining misconceptions has merit because students' prior knowledge predicts student achievement and learning, (e.g., Morrison & Lederman, 2003; Thompson & Zamboanga, 2003, 2004). Although prior knowledge may facilitate new learning (Beier & Ackerman, 2005; Chiesi, Spillich, & Voss, 1979; Hambrick & Engle, 2002), it can also impede learning (Lipson, 1982). Of particular relevance to research on psychological misconceptions and conceptual change are studies in which inaccurate knowledge interferes with conceptual change in science learning (Chinn & Malhotra, 2002; Dunbar, Fugelsand, & Stein, 2007).

Another view of misconceptions associates them with naive science and assumes that people who lack scientific knowledge hold theory-like beliefs that stand in for scientific explanations. This naive science view maintains that people commonly apply their intuitive notions to physics (McCloskey, 1983), astronomy (Miller & Brewer, 2010), biological evolution (Bloom & Weisberg, 2007), and other sciences.

According to a naive science view, students enter the field with many psychological misconceptions because they are ''naive psychologists,'' already endorsing many commonsense notions of mind and behavior at odds with scientific psychology. People seem prone to explain their own and others' actions in terms of mental states, intentions, and other psychological causes with or without instruction in psychological science. Indeed, developmental psychologists argue that a normal phase of human development is the acquisition of a theory of mind (Gelman & Legare, 2011). Others contend that the human mind contains a module dedicated to implicit understandings of the mind and behavior that may interact with the acquisition of scientifically based theories of mind (Geary, 2008). One example is the natural tendency for people to be mind–body dualists (Bloom, 2004), an ontological stance associated with (a) the tendency to hold dualistic misconceptions, such as the belief that in an out-of-body experience (OBE) the mind leaves the body (Bensley, 2003) and (b) belief in the reality of paranormal events (Alcock, 1987).

In contrast, we suspect that relatively few items in psychological misconceptions tests are answered incorrectly because students possess coherent, naive theories at odds with psychological science. Rather, we contend misconceptions usually stem from students accessing isolated, incorrect bits of knowledge acquired through informal, nonscientific sources, including everyday conversation, the media, fiction, rumors, and conversations (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012). The result is a set of commonsense concepts that often do not logically cohere (Thagard, 2012) and that diverge from psychological science. As such, psychological misconceptions often represent specific bits of knowledge at odds with specific psychological facts (Hughes, Lyddy, & Lambe, 2013; Taylor & Kowalski, 2014). Likewise, the status of theory of mind as ''theory'' in this technical sense has been seriously questioned (Ratcliffe, 2007). This makes psychological misconceptions seem more consistent with the notion of ''knowledge in pieces'' (diSessa, 2006) than as a part of naive psychological theories.

Psychological misconceptions are commonly assumed to be prevalent (Gardner & Brown, 2013; Taylor & Kowalski, 2014). However, there are considerable differences in the overall frequencies of misconceptions reported, with the percentages across studies reviewed in Lilienfeld, Lynn, Ruscio, and Beyerstein (2010) ranging from 28% to 71% (Hughes, Lyddy, & Lambe, 2013). These findings raise the possibility that there may be substantial differences in the endorsement of misconceptions across items. This raises the following important question: Are low-frequency false beliefs still misconceptions?

We propose that false, commonsense beliefs about the mind, brain, and behavior must logically be considered to be psychological misconceptions, regardless of their frequency. At the same time, this conceptual definition does not address the practical problem of the scientist who must demonstrate that a given false, commonsense belief actually exists. For the psychologist, this often requires demonstrating that a response occurs frequently enough to not be attributable to chance or guessing.

We define psychological misconceptions as false, commonsense beliefs because they typically originate from informal sources and are commonly held but are inconsistent with the consensus of research in psychology. The word ''commonsense'' also implies they are popular or prevalent, and they exist as part of the received knowledge of mind, brain, and behavior that people acquire through immersion in their culture. Often, the source of a misconception is unknown, yet they may seem familiar, intuitively true, and what ''everyone knows.'' Indeed, compared with other students, students who endorse more misconceptions tend to adopt a more intuitive and less scientific approach to knowledge claims and are more accepting of pseudoscientific and paranormal claims (Bensley, Lilienfeld, & Powell, 2014).

For the psychologist assessing the frequency of these false, commonsense beliefs in a group of people, the obvious question concerns what is meant by ''frequent.'' One criterion used by some authors (e.g., Brown, 1983; Taylor & Kowalski, 2012; Vaughan, 1977) is that a misconception occurs when at least 50% of respondents select an incorrect option. In the next section, we argue that problems with this criterion, especially in relation to certain ways of measuring misconceptions, can lead to faulty frequency estimates. More fundamentally, most tests of psychological misconceptions are characterized by psychometric shortcomings that make the accurate estimation of the frequencies of misconceptions problematic.

Shortcomings of Misconceptions Tests

One shortcoming of many misconceptions tests is the use of a true–false (T/F) response format (e.g., Brown, 1983; Gardner & Dalsing, 1986; Griggs & Ransdell, 1987; Kuhle, Barber, & Bristol, 2009; Lamal, 1979; Taylor & Kowalski, 2004; Vaughan, 1977). Consistently keying answers as either T or F can induce a response set and may help respondents detect the pattern of correct answers and guess correctly. Another problem with the usual T/F format is that it constrains responses to be completely true or completely false. However,

the inductive and informal reasoning used to build scientific theories is defeasible, often permitting conclusions that are only tentative, qualified, and provisional (Bensley, 1998).

Ambiguity in the expression of some misconception items has also made it difficult to conclude that an item is completely false (Brown, 1984) with qualifiers needed for items that are too ambiguous (Ruble, 1986). Hughes, Lyddy, and Kaplan (2013) found that ambiguously phrased items yielded higher levels of misconception endorsement than clearly phrased items.

Another shortcoming of previous misconception tests has been that respondents have often not been allowed to indicate they did not know an answer. One concern is that when respondents guess, this will produce an inaccurate estimate of their knowledge. In an attempt to eliminate the contribution of guessing, Gardner and Dalsing (1986) administered a 60-item version of the Test of Common Beliefs (TCB) of Vaughan (1977) in T/F format but added a third option of "don't know/no opinion." They found that students chose this option 12.2% of the time. After discarding such responses, the frequency of misconceptions was reduced by 8% on 14 common items. Adopting this strategy may control for guessing, but it also produces total test scores that are based on an unequal number of responses to items. Moreover, judging that one does not know an answer or has no opinion is not necessarily equivalent to the more continuously varying judgment of one's ability to provide a correct answer. Another problem is that responding with "no opinion" about a question might be an easy way to opt out of answering for respondents who lack motivation to answer the question but who may not lack relevant knowledge.

A better strategy may be to assess students' confidence and guessing as a separate metacognitive dimension from their knowledge of psychology by having them judge their certainty in the correctness of their answers. In one of the few studies to adopt this approach, Landau and Bavaria (2003) asked respondents to rate their confidence on a 5-point Likert-type scale after answering each of a test's 10 misconception questions. They found that respondents were significantly more confident on six of the incorrect items (misconceptions) than on correct responses; however, on 1 other item, respondents were significantly more confident of their correct answer than the incorrect answer. Landau and Bavaria concluded that, in general, students are more confident regarding misconceptions than regarding accurate beliefs.

Still, methodological limitations of their study preclude clear-cut interpretation of these findings. The authors tested only 10 misconceptions using a T/F format in which T statements were always misconceptions, raising the possibility of response bias. Also, they did not control for the number of *t*-tests they conducted, raising the possibility of Type I error. Had a conservative Bonferroni procedure had been applied, none of their comparisons would have been significant.

Taylor and Kowalski (2004) asked students to rate their confidence in their responses after answering each of their 48 test items (36 T/F items with F always correct and 12 filler items in which the correct response was T). To examine guessing, they collapsed across the four lower values of their 10-point confidence rating scale and treated these values as indicating guessing. The guessing items accounted for 28% of the responses. When they compared the accuracy of all the responses with the set of responses that did not include guessing, accuracy dropped slightly from 38.5% to 35.6%. Taylor and Kowalski also compared the mean confidence of all correct responses with the mean of all incorrect responses, finding no significant difference.

The lack of attention to the confidence of students in their endorsement of misconceptions leaves many questions unanswered. For example, it would be wrong to assume that a person endorsing a misconception is intentionally holding the incorrect belief. It is more likely that the person does not realize that the belief is false and instead tacitly accepts the misconception based on commonsense knowledge without considering an alternative view. This remains an important question because if students are generally unaware of which answers are misconceptions, they will not accurately perceive which aspects of their knowledge need correction. If they are more confident of their misconceptions than of their correct beliefs, they will presumably not be motivated to revise these misconceptions. Burton (2005) warned about the problems of estimating frequencies in T/F tests when test takers guess. Without assessing a student's tendency to guess, it is difficult to evaluate the probability of their answering correctly on T/F tests.

In a study attempting to deal with the problems of tests with T/F format in which misconceptions are always associated with a true response, Brown (1983) assembled a pool of multiple-choice, misconception questions. He established that any item be classed as misconception if at least 50% of the students selected an incorrect option. This is a fairly stringent criterion because the probability of selecting any response alternative should be 25% by chance alone. From this pool, he created 37 T/F items, wording 18 of the 37 as "false." Retaining the initial 50% criterion for the new T/F items, he found that only 19 of both true and false items were missed by at least 50% of the students, leading him to conclude that misconceptions may be less frequent than supposed. Nevertheless, keeping the same criterion for the T/F items may have led to overestimating misconceptions because if respondents were merely guessing, the expected frequency of choosing T or F would be .50, assuming a priori that T and F are equally probable. Consequently, if students were guessing, we would expect them to get an item wrong 50% of the time over the long run, and the average frequency of misconceptions on such a test would be 50%. Any response selected by nearly 50% of respondents would not differ significantly from chance.

Other studies have attempted to counter problems with the T/F format, by having half of correct items keyed false and intermixing misconceptions with more conventional general psychology questions (Kowalski & Taylor, 2009) and by using a multiple-choice format (McCutcheon, 1991). Nevertheless, it is not known how well these alternate formats solved the problems of T/F tests.

Addressing the criticism that some misconceptions tests contain out-of-date items, four recent studies have used the book, *50 Great Myths of Popular Psychology* by Lilienfeld et al. (2010) as a source for developing misconception items (Furnham & Hughes, 2014; Gardner & Brown, 2013; Hughes, Lyddy, & Kaplan, 2013; Taylor & Kowalski, 2012). The book discusses many commonsense misconceptions across many domains of psychology and presents research documenting their prevalence and refuting the misconceptions. Using this book as a starting point, Gardner and Brown developed a test to examine the contribution of the T/F test format. They constructed some misconceptions in T format and others in F format to examine the effect of wording the truth value of items. Also addressing the problem of misconceptions not being completely false, they used a Likert-type scale ranging from *completely false* to *completely true*. As in the study by Gardner and Dalsing (1986), Gardner and Brown allowed respondents to report they did not know the answer, using the *don't know/no opinion* option.

Their new test showed good internal consistency, and scores were not significantly affected by the direction of item keying. In general, students endorsed misconceptions near the midpoint of their scale, leading the authors to conclude that students judged misconceptions to be partly false and partly true. Although Gardner and Brown's (2013) continuous response format does not show how much knowledge students possessed or the frequency of misconceptions, the authors should be commended for their attention to the psychometrics of their instrument. Although some researchers have provided evidence of the psychometrics of their tests (e.g., McCutcheon, 1991; Vaughan, 1977), most studies reporting frequencies of misconceptions have not done so. Next, we describe our efforts to develop a new psychometrically sound measure of psychological misconceptions that used the Lilienfeld et al.'s (2010) book as the source of items but was designed to avoid the shortcomings of previous tests.

A New Test Addressing the Shortcomings

To help solve problems associated with the T/F format, we constructed the TOPKAM to have a forced-choice format in which a response option reflecting accurate psychological knowledge is counterposed against the corresponding psychological misconception based on literature reviews in Lilienfeld et al. (2010) and other sources. Testing whether people endorse misconceptions when a better alternative is available can distinguish tacit acceptance from the failure to consider the alternative.

Supporting the utility of the new forced-choice format of the TOPKAM, Taylor and Kowalski (2012) used and adapted several items from the TOPKAM (Bensley & Lilienfeld, 2010) and compared the number correct on this modified version to a T/F item format test with similar questions. They found that introductory psychology students scored 41.3% correct on the modified TOPKAM, whereas the same students scored 33.1% correct on the T/F test. Taylor and Kowalski judged the forced-choice format to be better than the T/F version, although they did not test students' confidence in their responses. To address the fact that psychological knowledge is tentative and rarely completely true or false, TOPKAM's instructions ask test takers to answer questions by judging which option is "best" and question stems ask them to select the option that is "most true."

To address the problem of guessing and the unequal number of responses when don't know/no opinion leads to eliminating responses, the TOPKAM treats guessing as part of a separate certainty dimension. Specifically, respondents rate the certainty of the correctness of their answer following each question, thereby postdicting the accuracy of their answer. Postdiction involves providing an immediate, retrospective measure of metacognitive monitoring (Hacker, Bol, & Keener, 2010).

A previous study revealed that the TOPKAM showed adequate test–retest reliability and internal consistency. In the present study, we examined the frequencies of individual TOPKAM items to determine whether they occurred at consistently high enough frequencies across two samples to be considered psychological misconceptions. We expected that if the TOPKAM reliably measured misconceptions, it should show consistency in endorsement of the same items across the two samples. If item-by-item consistency were found across the two samples, this finding would provide grounds for the further identification of specific misconceptions, applying the stringent criterion that an item would be considered a misconception if it consistently exhibited a frequency that significantly exceeded 50%, the level expected by chance. Based on our findings that higher misconceptions scores on the TOPKAM were associated with more intuitive thinking (Bensley et al., 2014), we expected that participants would be significantly more certain of their responses to misconception items frequently answered incorrectly than to items infrequently answered incorrectly.

Method

Participants

As part of psychology departmental assessment in the 2010–2011 and 2011–2012 academic years, we tested two samples of psychology students at a small, mid-Atlantic comprehensive university. Both the 2010–2011 sample with 205 students and the 2011–2012 sample with 191 students included students enrolled in two beginning courses and in senior capstone courses required for majors, and the 2010–2011 sample also included beginning and finishing graduate counseling psychology students. We eliminated data from seniors in the beginning courses who were beginning the major because our learning outcome assessment procedures routinely assess them later in senior capstone courses.

We retained data from 162 participants from the 2010–2012 sample and 173 from the 2011–2012 sample who completed all TOPKAM items and corresponding certainty ratings and who

Table 1. Description of Two Samples Administered the TOPKAM in Consecutive Academic Years.

	2010–2011 Sample	2011–2012 Sample
N	162	173
Male	22.2%	22.9%
Female	77.8%	77.1%
First year	5%	11.8%
Sophomore	28.6%	30.6%
Junior	16.8%	23.5%
Senior	30.4%	34.1%
Graduate	19.3%	0%
Age	$M = 22.20$, $SD = 4.49$	$M = 20.61$, $SD = 3.33$
White non-Hispanic	76.3%	66.1%
African American	20.6%	28.5%
Hispanic	1%	2.4%
Other ethnic	2.1%	3.0%

Note. TOPKAM = Test of Psychological Knowledge and Misconceptions.

consented to let us use their data. Missing any misconception item or certainty rating for that item would not allow us to accurately estimate the frequency of individual items in relation to certainty ratings. In the 2010–2011 administration, we placed the TOPKAM toward the end of the booklet of forms to be completed. Because some participants did not follow the directions to complete certainty ratings after completing each item, we discarded the data from those 24 participants in the 2010–2011 sample. For the second sample, we administered the TOPKAM first and read the TOPKAM instructions to students resulting in the exclusion of only nine students who did not complete their certainty ratings. We also excluded an additional 12 students from the 2011–2012 sample who had previously received critical thinking instruction related to the study or who indicated they had read at least part of the Lilienfeld et al.'s (2010) book before the first day of class.

Table 1 shows demographic characteristics of the 2010–2011 in the middle column and the same characteristics in the right-hand column for the 2011–2012 sample. Inspection of Table 1 shows the two samples had similar proportions of these characteristics, except only the first sample contained graduate students.

Measures

The TOPKAM contains 40 questions presented in a forced-choice, two-response format for assessing both factual knowledge of psychology and susceptibility to common psychological misconceptions (Bensley & Lilienfeld, 2010). We selected 40 of the misconceptions discussed in the 50 essays in Lilienfeld et al. (2010), selecting at least 2 items from each of the 11 general categories of psychological myths in the book to promote a representative sampling of misconception items. The domains comprised myths about the brain and perception; development and aging; memory, intelligence, and learning; consciousness, emotion, and motivation; interpersonal behavior, personality, mental illness, psychology, and the law; and psychological treatments. Scores on the TOPKAM can range from 0 to 40, with higher scores reflecting more accurate knowledge of psychology or alternately range from 0 to 40 incorrect, scored for psychological misconceptions.

Items described misconceptions in language that combined everyday language with scientific language while avoiding jargon. For example, one question asked "Which is most true about the Rorschach (inkblot) Test?" followed by two response options: "(a) It is like a 'psychological x-ray' because it can penetrate the unconscious mind and tell a great deal about personality" versus "(b) It can detect marked thinking disturbances but is not effective in detecting depression or anxiety disorders."

Instructions asked respondents first to select the "best" and "most true" answer for each question and then to rate their certainty of the correctness of their answer, using a 5-point Likert-type scale ranging from 1 (*not at all certain*) to 5 (*completely certain*) that appeared at the end of instructions and before the first item. Instructions asked students to answer each test question by bubbling in their answer on a Scantron form and then to write the certainty rating next to the number of the question after answering it.

Bensley et al. (2014) found that the internal consistency of the TOPKAM across the entire sample was $KR20 = .73$, suggesting modest but adequate reliability. Test–retest reliability at 4 weeks was $r = .69$. Supporting its validity, multiple regression analyses revealed that SAT, psychology grade point average, measures of critical thinking skill and dispositions, and measures of the ability to distinguish pseudosciences from sciences and poorly supported practices from well-supported practices, all significantly predicted the number correct on the TOPKAM (Bensley et al., 2014).

To assess test-taking motivation, students in both samples completed the Student Opinion Scale (SOS), a 10-item, self-report measure of assessment motivation. The SOS is reported on a scale with five different categorical response options ranging from A (*strongly disagree*) to E (*strongly agree*). The SOS yields two factors, effort expended and assessment importance, with internal consistencies ranging from .80 to .89 in a sample of over 15,000 students (Sundre & Moore, 2002).

Students in both samples completed a demographics form that assessed their academic background, class rank, gender, and ethnicity. The two samples also completed measures to assess their critical thinking dispositions. Results from these analyses are reported elsewhere (Bensley, Rainey, Lilienfeld, & Kuehne, in press). The assessment forms were assembled into booklets that varied in the 2 years. The consent form appeared on top both years but was followed by different measures in the 2 academic years. In 2010–2011, the first forms were disposition measures and the demographics form, followed by the TOPKAM and as mentioned before, the TOPKAM appeared first after the consent form in 2011–2012.

Procedure

During their regularly scheduled classes, the first author administered the TOPKAM and other measures to beginning majors

early in the fall semester and to senior majors, typically in the middle of the fall and spring semesters. He distributed the booklets of assessment forms along with a Scantron form for writing their TOPKAM answers. He discussed the consent form and instructed all students to carefully read the instructions, do their best, report honestly, and complete forms in the order in which they were presented. He assured students that their data would be treated confidentially, and all students received course credit for their participation except for senior internship students in both samples and graduate students in the first sample. In 2011–2012, after explaining the consent form, he read the directions to the TOPKAM to participants, emphasizing that they should rate the certainty of their answer after completing each individual item by writing their rating next to their answer on the Scantron form.

Results

We compared the TOPKAM scores of the 24 students in the 2010–2011 sample who did not complete the certainty ratings with the 162 students who did and found no significant difference in the scores. To examine the level of test-taking motivation, we calculated descriptive statistics on the SOS completed at the end of each assessment session. Respondents' total scores on the SOS in 2010–2011 ($M = 37.18$, $SD = 6.99$) revealed moderate levels of motivation comparable to the levels in the 2011–2012 sample ($M = 36.13$, $SD = 7.13$).

Frequency Analyses of Items

We conducted frequency analyses on the 40 items in each administration of the TOPKAM and identified those questions answered incorrectly by at least 50% in each sample. Table 2 shows the 18 items answered incorrectly by at least 50% of participants in both samples. One other item, not reported in Table 2, TOPKAM Question 9 which states that to overcome a mental disorder a person must confront the root causes from childhood, was answered incorrectly by 55.1% in the 2011–2012 sample but only by 41.4% in the 2010–2011 sample.

We conducted χ^2 tests on each of the 18 items missed by at least 50% in both samples separately for each academic year to test whether each item's frequency differed significantly from the expected frequency of 50% for that year. To maintain $\alpha = .05$ for the 18 χ^2 tests conducted for each respective year, we used a conservative Bonferroni correction procedure. Table 2 shows that 12 of the same items exhibited frequencies significantly higher than 50%. Three other items displayed frequencies that were significantly greater than 50% in 1 year. The item referring to dyslexia's defining feature as reversing letters was significant only in 2010–2011, whereas the items stating that hypnosis is a special state of consciousness and that positive attitudes stave off cancer were significant only in the 2011–2012 sample.

Table 3 shows that 21 of the same items had frequencies that were less than 50% in both samples. Next, we conducted two sets of similar χ^2 tests on each of the 21 items answered incorrectly by less than 50% in both samples, separately for each academic year, to test whether each item's frequency differed significantly from the expected frequency of 50% for that year, again using a Bonferroni correction. Table 3 shows that 15 of the same items had frequencies that were significantly lower than 50% in both samples. Three other items were significantly lower than 50% only in 2010–2011, whereas no other item was significant only in 2011–2012. The items referring to the validity of graphology (handwriting analysis) and the validity of the Rorschach inkblot test and the item stating that people behave abnormally during a full moon were all missed significantly less than 50% by those in the 2010–2011 sample, which contained graduate counseling students.

Next, we examined the consistency of the pairs of items for the two sample years displayed in Table 2 and overall.

Table 2. Percentage Incorrect for 18 TOPKAM Items Missed by at least 50% in Two Samples.

Item	Summary Description of Misconception	2010–2011	2011–2012
1	Shock therapy is dangerous, brutal treatment	58.0	54.9
2	Dyslexia's defining feature is reversing letters	62.3*	60.1
4	Better to vent your anger or "blow off steam"	53.7	55.5
5	Most people only use 10% of their brains	51.2	60.7
12	Hypnosis is a special state of consciousness	51.9	65.3*
13	Children's personality very similar to parents	82.7*	86.1*
15	Amnesiacs cannot remember previous life	81.5*	88.4*
16	Expert profilers much more accurate than other people	67.9*	68.8*
18	Individuals repress memory of traumatic experiences	76.5*	79.8*
20	People with severe mental illness prone to violence	66.0*	75.1*
21	Children of alcoholics have lower self-esteem	90.7*	93.1*
23	Raising self-esteem improves academics	90.1*	96.0*
24	Stick with initial hunches about test answers	86.4*	90.2*
30	Positive attitudes stave off cancer	52.5	65.3*
35	Should match teaching style to learning style	87.7*	90.2*
36	Subliminal messages induce people to consume	64.2*	72.8*
37	Adolescence is time of psychological turmoil	74.1*	69.4*
40	Some people right-brained, some left-brained	67.3*	67.6*

Note. TOPKAM = Test of Psychological Knowledge and Misconceptions from Bensley and Lilienfeld (2010).
*$p < .05$ after significant χ^2 test and application of the conservative Bonferroni procedure to control for multiple significance tests conducted.

Table 3. Percentage Incorrect for 21 TOPKAM Items Missed by Less Than 50% in Two Samples.

Item	Summary Description of Misconception	2010–2011	2011–2012
3	Traits with high heritability cannot be modified	26.5*	30.6*
6	Clinician's using intuition make better decisions	17.3*	22.0*
7	People more romantically attracted to people like them	29.6*	34.7*
8	Memory contains a perfect record of all our experiences	9.3*	13.9*
10	Schizophrenics have split personalities	39.5	49.7
11	There has been epidemic of genuine cases of autism	19.8*	21.4*
14	Lie detector almost never misidentifies someone	12.3*	13.3*
17	People with enduring happiness received many positive	19.1*	30.1*
19	People in "middle years" experience a mid-life crisis	40.7	41.6
22	People can learn a new language while asleep	32.1*	37.6*
25	Graphology is as accurate as personality inventories	30.9*	43.4
26	Rorschach inkblot test can penetrate unconscious mind	30.9*	41.6
27	When see, tiny rays or light particles emitted from eyes	11.7*	18.5*
28	Older people are more lonely, depressed, and cranky	31.5*	36.4*
29	Dreams often have important symbolic meaning	25.9*	30.6*
31	If groups score differently on IQ test, then it is biased	28.4*	16.8*
32	Men and women very different in how communicate	48.1	49.7
33	People behave abnormally during full moon	30.2*	41.6
34	People claiming have ESP not better predicting future	15.4*	21.4*
38	Consciousness really leaves in out-of-body experience	24.7*	23.1*
39	Larger group more likely to help person in an emergency	24.7*	22.5*

Note. TOPKAM = Test of Psychological Knowledge and Misconceptions from Bensley and Lilienfeld (2010).
*$p < .05$ after significant χ^2 test and application of the conservative Bonferroni procedure to control for multiple significance tests conducted.

Inspection of the frequency scores for individual items of all 40 items in the two sample years revealed that the frequencies in each pair were similar, ranging in difference from 0.3% to 13.4%. To analyze the apparently similar frequencies in the respective item pairs in the two sample years, we calculated the correlation between the frequencies of the 40 pairs of items, using the frequencies in the 2 years for each item as the unit of analysis. The Spearman rank-order correlation across all 40 items, $r_s = .97$, $p < .001$, was significant and very high. Likewise, the Spearman rank-order correlation on the pairs of frequencies for the 18 items that were answered incorrectly by at least 50% in both samples, $r_s = .91$, $p < .001$, was very high. The correlation between the pairs of frequencies of the 12 items that were significantly higher than 50%, $r_s = .88$, $p < .001$, was also high. These results demonstrate substantial consistency in the frequency of endorsement across the two samples.

Analyses of Certainty Ratings

To examine the metacognitive monitoring of student' knowledge of psychology, we conducted analyses of their certainty ratings accompanying each TOPKAM question. To test whether respondents were more confident of their incorrect responses than their correct ones, as found by Landau and Bavaria (2003), we compared the certainty ratings of respondents when they answered correctly with those of respondents answering incorrectly. We conducted 40 independent samples *t*-tests on the certainty of correct versus incorrect responses for each of the items in the 2010–2011 and the 2011–2012 samples. Following Landau and Bavaria (2003), we first report the probabilities of these comparisons unadjusted. Of the 40 comparisons in 2010–2011, 29 were not significant. Six of the 40 showed significantly higher certainty ratings for correct responses to items than for incorrect, whereas for 5 other items, the incorrect responses showed significantly higher certainty ratings. To maintain $\alpha = .05$ for these 40 comparisons, we applied the Bonferroni procedure and found that Question 6, which refers to clinicians' use of intuition resulting in better decisions, showed significantly higher certainty for the correct response; whereas Question 23 which refers to the claim that raising self-esteem improves academic performance, the incorrect response showed significantly higher certainty.

Next, we conducted a similar set of 40 *t*-tests on the certainty of correct versus incorrect items from 2011 to 2012. We first report the probabilities of these comparisons unadjusted. Of the 40 comparisons in 2011–2012, 22 were not significant and 18 were significant. Table 4 shows that 6 of the 40 showed significantly higher certainty ratings for correct responses to items than for incorrect while for 12 other items, the incorrect responses showed significantly higher certainty. Applying the Bonferroni procedure, we found only 4 items passed this stringent test of significance. Question 27, which refers to visual perception involving emission of light from the eye, and Question 28, which concerns elderly people being especially lonely and depressed, showed significantly higher certainty for the correct response. In contrast, Question 5, which concerns the idea that people only use 10% of the brain, and Question 24, which concerns always sticking with one's initial answer on a test, the incorrect response showed significantly higher certainty.

The initial *t*-test results show agreement on certainty in both years on only 8 items with 3 showing greater certainly for correct responses and 5 showing greater certainty for incorrect responses. After applying the Bonferroni correction, none of

Table 4. Significance Tests of Correct Versus Incorrect Certainty on TOPKAM Items in Two Samples.

Item	Summary Description of Misconception	2010–2011	2011–2012
5	Most people only use 10% of their brains	Incorrect	Incorrect*
6	Clinician's using intuition make better decisions	Correct*	NS
7	People more romantically attracted to people like them	Correct	Correct
8	Memory contains a perfect record of all our experiences	NS	Correct
9	Focus on origin of mental problem to recover from it	Correct	NS
10	Schizophrenics have split personalities	NS	Incorrect
13	Children's personality very similar to parents	NS	Incorrect
14	Lie detector almost never misidentifies someone	Correct	NS
15	Amnesiacs cannot remember previous life	NS	Incorrect
16	Expert profilers much more accurate than other people	NS	Incorrect
18	Individuals repress memory of traumatic experiences	Incorrect	Incorrect
21	Children of alcoholics have lower self-esteem	Incorrect	Incorrect
23	Raising self-esteem improves academics	Incorrect*	Incorrect
24	Stick with initial hunches about test answers	NS	Incorrect*
25	Graphology is as accurate as personality inventories	NS	Incorrect
27	When see, tiny rays or light particles emitted from eyes	Correct	Correct*
28	Older people are more lonely, depressed, and cranky	Correct	Correct*
34	People claiming have ESP not better predicting future	NS	Correct
35	Should match teaching style to learning style	Incorrect	Incorrect
36	Subliminal messages induce people to consume	NS	Incorrect
39	Larger group more likely to help person in an emergency	NS	Correct

Note. All items not listed in the table are not significant; "correct" refers to significantly greater mean certainty of correct responses; "incorrect" refers to significantly greater mean certainty of incorrect responses. NS = not significant before the Bonferroni procedure was applied.
*$p < .05$ after conservative Bonferroni procedure was applied.

the comparisons was significant for an item across both years, and only six were significant in either year. When we applied the appropriate correction for Type I error, our results provided little, consistent support for the hypothesis that people are more certain (confident) of their incorrect responses than correct ones.

Nevertheless, taking into account the frequency of response for individual items might provide a more sensitive test of the hypothesis. Accordingly, we conducted additional tests that used our earlier findings from the 2010–2011 sample to compare the certainty ratings of all items answered incorrectly by significantly more than 50% shown in Table 2 to the certainty ratings of those answered incorrectly by significantly less than 50% of participants shown in Table 3. A paired samples *t*-test, $t(161) = 5.01$, $p < .001$, $d = .39$, showed that the mean certainty of the items answered incorrectly the most ($M = 3.73$, $SD = .60$) was significantly higher than the mean certainty of the items answered incorrectly the least ($M = 3.54$, $SD = .59$ and associated with a small to medium effect size.

Repeating this procedure, we compared the mean certainty for the items in 2011–2012 sample that were answered incorrectly by significantly more to the mean certainty answered incorrectly by significantly less than 50% of participants. A paired samples *t*-test, $t(172) = 7.84$, $p < .001$, $d = .60$, showed that the mean certainty of the items more often answered incorrectly ($M = 3.68$, $SD = .58$) was significantly greater than the mean certainty of the items incorrectly answered less often ($M = 3.38$, $SD = .66$) and associated with a medium effect size.

Discussion

We defined psychological misconceptions as commonsense beliefs about the mind, brain, and behavior that are held contrary to what is known from psychological research and argued for the importance of evaluating the frequency of individual misconceptions. We described the development of the TOPKAM, a measure designed to address shortcomings of previous tests. We used a forced-choice format to avoid potential response biases that can arise in T/F instruments in which true responses are always scored as misconceptions. Identification of misconceptions and the correct responses for TOPKAM items were based on recent literature reviews on psychological misconceptions (e.g., Lilienfeld et al., 2010). To address the problem of the inherently provisional nature of most psychological knowledge, TOPKAM questions ask participants to choose which of the two response options was most true based on their knowledge of psychology.

We tested whether individual misconception items were endorsed at levels significantly greater than 50% based on our assumption that psychological misconceptions should be demonstrated to occur at frequencies higher than what would be expected by chance. Our conservative analyses of the frequency data identified 12 of the 40 misconception items from the TOPKAM that were endorsed by significantly more than 50% of students in both samples. Our results suggest that psychological misconceptions identified by these stringent criteria may not be as common as some have assumed (e.g., Gardner & Brown, 2013). Inspection of Tables 2 and 3 show a wide range of frequencies in the endorsement of misconceptions for different items. Nevertheless, the frequencies of endorsement of the individual misconceptions across two samples were highly consistent.

We found that high-frequency misconception items could be distinguished from low-frequency misconception items based

on respondents' ratings of the certainty of their answers. Students in both samples provided significantly higher certainty ratings on items more often answered incorrectly (significantly more than 50%) than on items less often answered incorrectly (significantly less than 50%). These results support the hypothesis that students are more confident of misconception items they more often answer incorrectly than of those they less often answer incorrectly. That being said, the effect size was small in the first sample and medium in the second sample, with only fairly small fractions of a point difference in mean certainty between more frequently and less frequently endorsed misconception items in both samples.

More generally, these results support the hypothesis of a metacognitive deficit regarding misconceptions in which students show a lack of awareness of their lack of knowledge. In particular, our results suggest that students tend to be overconfident of their answers on the misconception items most frequently answered incorrectly. When students know the least, they tend to be most certain that they are right. In the language of metacognition, their certainty judgments are not well calibrated with their performance. Further evidence of overconfidence comes from a follow-up study in which we found that students significantly overestimated their TOPKAM scores when asked to postdict them (Bensley et al., in press).

The results also support the utility of operationalizing psychological misconceptions for research purposes as high-frequency errors. When frequency was not taken into account and we compared the certainty of all items answered incorrectly with those answered correctly, we found no significant difference. This failure to replicate the findings of Landau and Bavaria (2003) is consistent with a conservative interpretation of their results because with a correction for a Type I error their positive results disappear. Our results suggest that further study of the relations between metacognitive monitoring and the TOPKAM is warranted.

It is useful to take stock of the status of this knowledge acquired through the application of our test. We believe that the encouraging data supporting the reliability and validity of the TOPKAM and our conservative approach to identifying misconceptions may facilitate the identification of specific psychological misconceptions and a better understanding of how they relate to other academic and instructional variables. As noted by Keil (2012), more research is needed on the relations between people's commonsense psychological concepts and the acquisition of well-supported theories of psychological science. Although the TOPKAM seems to be a useful instrument to operationalize the testing of misconceptions, future studies should combine it with other reliable and valid measures to provide convergent meaning to the construct of psychological misconceptions (Grace, 2001).

Identifying misconceptions may also be useful to teachers of psychology who have limited class time. If instructors obtained data like that found in Tables 2 and 4, they could combine this information to select misconceptions of which students are especially confident and most in need of correction. For example, instructors might decide to cover a misconception such as Item 23, which refers to the claim that raising self-esteem improves academic performance, because Table 2 shows that over 90% of respondents accepted this false idea and Table 4 showed they were consistently overconfident of it (see Taylor & Kowalski, 2012, for similar results).

The results in Tables 3 and 4 may also help to identify mistaken ideas that require less attention, given that other instructors obtain similar data with the same items. For example, students were quite accurate and appropriately confident in answering Item 27, regarding the idea that the eyes emit tiny rays, with only 11.7% answering the question incorrectly in the first sample and 18.5% in the second sample. Although Winer, Cottrell, Gregg, Fournier, and Bica (2002) found that college students often answered questions indicating they believed that the eyes emitted light in vision, more recent testing with a question similar to the TOPKAM item by Taylor and Kowalski (2012) showed that only 15% answered incorrectly. This percentage is very similar to the 15.9% who answered it incorrectly when we reanalyzed that TOPKAM item on our beginning majors who had taken introductory psychology. These low frequencies from different samples suggest that this false belief may be less pressing to address than some others in introductory psychology courses.

At the same time, good arguments could be mustered for why all of the TOPKAM misconceptions, not just high frequency ones, should be addressed in instruction and used in learning outcomes assessment. Although the Bonferroni test was useful in differentiating high- from low-frequency items in terms of certainty, it is a very conservative test that may have led us to overlook some misconceptions only slightly above 50%.

Perhaps a stronger objection to calling only high frequency, incorrect answers "misconceptions," is that psychological misconceptions naturally occur at different, and sometimes lower, frequencies because they are of different types and from different sources. A case in point is the misconception that the mind actually leaves the body in the OBE, which was endorsed by 24.7% in the first sample and 23.1% in the second. These frequencies are strikingly similar to the 25% rate of incidence of OBE reported by college students across many studies (Alvarado, 2000). Blanchfield, Bensley, Hierstetter, Mahdavi, and Rowan (2007) found that 24.3% and 28.4% of their college student sample reported having had an OBE. Those who had an OBE were significantly more likely to believe that the mind actually leaves the body and to endorse mind–body dualism than those not having had one. Consequently, the OBE question may be responded to differently depending on a student's ontological stance (Hughes, Lyddy, & Lambe, 2013). More research is needed on possible differences between misconceptions as they relate to psychological and instructional variables. Different kinds of misconceptions are likely to require different instructional approaches (Lewandowsky et al., 2012).

Finally, perhaps some lower frequency items represent misunderstandings that are highly consequential even at lower frequencies. For example, incorrectly accepting the lower frequency misconception that the Rorschach inkblot test can

penetrate the unconscious mind and provide accurate diagnostic information regarding mood and anxiety disorders could lead many persons to accept a therapist's poorly supported conclusions.

Our preliminary results with the TOPKAM, if replicated in other samples, may offer a promising new way to study psychological misconceptions and reduce misconceptions. For example, our results suggest that instructors should take into account that students are often overconfident of their faulty beliefs, especially those that are frequently held. As a consequence, students may not question what they think they know. Future studies with the TOPKAM should investigate other ways that psychological misconceptions may differ from each other and further examine what students know about their knowledge of psychology.

Acknowledgments

We thank Lauren A. Powell for helping with data input and analysis, the instructors who allowed us to assess their students in various courses, and the students who consented to let us use their data. Thanks also to reviewers and the editor, Andrew Christopher, for their helpful comments.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

Alcock, J. E. (1987). Parapsychology: Science of the anomalous or search for the soul? *Behavioral and Brain Sciences*, *10*, 553–565.

Alverado, C. S. (2000). Out-of-body experiences. In E. Cardena, S. J. Lynn, & S. C. Krippner (Eds.), *Varieties of anomalous experience: Examining the scientific evidence* (pp. 183–218). Washington, DC: American Psychological Association.

Beier, M. E., & Ackerman, P. L. (2005). Age, ability and the role of prior knowledge on the acquisition of new domain knowledge: Promising results in a real-world learning environment. *Psychology and Aging*, *20*, 341–355.

Bensley, D. A. (1998). *Critical thinking in psychology: A unified skills approach*. Pacific Grove, CA: Brooks/Cole.

Bensley, D. A. (2003). Can minds leave bodies? A cognitive science perspective. *Skeptical Inquirer*, *27*, 34–39.

Bensley, D. A., & Lilienfeld, S. O. (2010). The test of psychological knowledge and misconceptions. Unpublished manuscript, Worth Publishers, New York, NY.

Bensley, D. A., Lilienfeld, S., & Powell, L. A. (2014). A new measure of psychological misconceptions: Relations with academic background, critical thinking, and acceptance of paranormal and pseudoscientific claims. *Learning and Individual Differences*, *36*, 9–18.

Bensley, D. A., Rainey, C., Lilienfeld, S. O., & Kuehne, S. (in press). What do psychology students know about what they know in psychology? *Scholarship of Teaching and Learning in Psychology*.

Best, J. B. (1982). Misconceptions about psychology among students who perform highly. *Psychological Reports*, *51*, 239–244.

Blanchfield, B., Bensley, D. A., Hierstetter, J., Mahdavi, A., & Rowan, V. (2007, March). *Reported OBE occurrence: Be careful how you ask*. Poster presented at the 78th Meeting of the Eastern Psychological Association, Philadelphia, PA.

Bloom, P. (2004). *Descartes' baby: How the science of child development explains what makes us human*. New York, NY: Basic Books.

Bloom, P., & Weisberg, D. S. (2007). Childhood origins of adult resistance to science. *Science*, *316*, 996–997.

Brown, L. T. (1983). Some more misconceptions about psychology among introductory psychology students. *Teaching of Psychology*, *10*, 207–210.

Brown, L. T. (1984). Misconceptions about psychology are not always what they seem. *Teaching of Psychology*, *11*, 75–78.

Burton, R. F. (2005). Multiple-choice and true/false tests: Myths and misapprehensions. *Assessment and Evaluation in Higher Education*, *30*, 65–72.

Chiesi, H. L., Spillich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, *18*, 257–273.

Chinn, C. A., & Malhotra, B. A. (2002). Children's responses to anomalous scientific data: How is conceptual change impeded? *Journal of Educational Psychology*, *94*, 327–343.

diSessa, A. A. (2006). A history of conceptual change research: Threads and fault lines. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 265–281). Cambridge, England: Cambridge University Press.

Dunbar, K. N., Fugelsang, J. A., & Stein, C. (2007). Do naïve theories ever go away? Using brain and behavior to understand changes in concepts. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 193–205). New York, NY: Lawrence Erlbaum.

Furnham, A., & Hughes, D. J. (2014). Myths and misconceptions in popular psychology: Comparing psychology students and the general public. *Teaching of Psychology*, *41*, 256–261.

Gardner, R. M., & Brown, D. L. (2013). A test of contemporary misconceptions in psychology. *Learning and Individual Differences*, *24*, 211–215.

Gardner, R. M., & Dalsing, S. (1986). Misconceptions about psychology among college students. *Teaching of Psychology*, *13*, 32–34.

Geary, D. C. (2008). An evolutionary informed education science. *Educational Psychologist*, *43*, 179–195.

Gelman, S. A., & Legare, C. H. (2011). Concepts and folk theories. *Annual Review of Anthropology*, *40*, 379–398.

Grace, R. C. (2001). On the failure of operationism. *Theory & Psychology*, *11*, 5–33.

Griggs, R. A., & Ransdell, S. E. (1987). Misconceptions tests or misconceived tests? *Teaching of Psychology*, *6*, 210–214.

Hacker, D. J., Bol, L., & Keener, M. C. (2010). Metacognition in education: A focus on calibration. In J. Dunlovsky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 429–455). New York, NY: Psychology Press.

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology*, *44*, 339–387.

Higbee, K. L., & Clay, S. L. (1998). College students' beliefs in the ten-percent myth. *The Journal of Psychology*, *132*, 469–476.

Hughes, S., Lyddy, F., & Kaplan, R. (2013). The impact of language and response format on student endorsement of psychological misconceptions. *Teaching of Psychology*, *40*, 31–37.

Hughes, S., Lyddy, F., & Lambe, S. (2013). Misconceptions about psychological science: A review. *Psychology learning and Teaching*, *12*, 20–31.

Keil, F. C. (2012). Does folk science develop? In S. M. Carver & J Shrager (Eds.), *The journey from child to scientist: Integrating cognitive development and the education sciences* (pp. 67–86). Washington, DC: American Psychological Association.

Kowalski, P., & Taylor, A. K. (2009). The effect of refuting misconceptions in the introductory psychology class. *Teaching of Psychology*, *36*, 153–159.

Kuhle, B. X., Barber, J. M., & Bristol, A. S. (2009). Predicting students' performance in introductory psychology from their psychology misconceptions. *Journal of Instructional Psychology*, *36*, 119–124.

Lamal, P. A. (1979). College students' common beliefs about psychology. *Teaching of Psychology*, *6*, 155–158.

Landau, J. D., & Bavaria, A. J. (2003). Does deliberate source monitoring reduce students' misconceptions about psychology? *Teaching of Psychology*, *30*, 311–314.

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*, 106–131.

Lilienfeld, S., Lynn, S., Ruscio, J., & Beyerstein, B. L. (2010). *50 great myths of popular psychology: Shattering widespread misconceptions about human behavior.* West Sussex, UK: Wiley-Blackwell.

Lipson, M. Y. (1982). Learning new information from text: The role of prior knowledge and reading ability. *Journal of Reading Behavior*, *14*, 243–261.

McCloskey, M. (1983). Native theories of motion. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 299–323). Hillsdale, NJ: Lawrence Erlbaum.

McCutcheon, L. E. (1991). A new test of misconceptions about psychology. *Psychological Reports*, *68*, 647–653.

McKeachie, W. J. (1960). Changes in scores on the northwestern misconceptions test in six elementary psychology courses. *Journal of Educational Psychology*, *51*, 240–244.

Miller, B. W., & Brewer, W. F. (2010). Misconceptions of astronomical distances. *International Journal of Science Education*, *32*, 1549–1560.

Morrison, J. A., & Lederman, N. G. (2003). Science teachers' diagnosis of understanding of students' preconceptions. *Science Education*, *87*, 849–867.

Piquette, J. S., & Heikkinen, H. W. (2005). Strategies reported used by instructors to address student alternate conceptions in chemical equilibrium. *Journal of Research in Science Teaching*, *42*, 1112–1134.

Ratcliffe, M. (2007). *Rethinking commonsense psychology: A critique of folk psychology, theory of mind, and simulation*. New York, NY: Palgrave.

Ruble, R. (1986). Ambiguous psychological misconceptions. *Teaching of Psychology*, *13*, 34–36.

Standing, L., & Huber, H. (2003). Do psychology courses reduce belief in psychological myths? *Social Behavior and Personality*, *31*, 585–591.

Sundre, D. L., & Moore, D. L. (2002). The student opinion scale: A measure of examinee motivation. *Assessment Update*, *14*, 8–9.

Taylor, A. K., & Kowalski, P. (2004). Naïve psychological science: The prevalence, strength, and sources of misconceptions. *The Psychological Record*, *54*, 15–25.

Taylor, A. K., & Kowalski, P. (2012). Students' misconceptions in psychology: How you ask matters . . . sometimes. *Journal of the Scholarship of Teaching and Learning*, *12*, 62–77.

Taylor, A. K., & Kowalski, P. (2014). Student misconceptions: Where do they come from and what can do? In V. A. Benassi, C. E. Overson, and C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum* (pp. 259–273). Washington, DC: American Psychological Association (Division 2).

Thagard, P. (2012). *The cognitive science of science*. Cambridge, MA: MIT Press.

Thompson, R. A., & Zamboanga, B. L. (2003). Prior knowledge and its relevance to student achievement in introduction to psychology. *Teaching of Psychology*, *30*, 96–101.

Thompson, R. A., & Zamboanga, B. L. (2004). Academic aptitude and prior knowledge as predictors of student achievement in introduction to psychology. *Journal of Educational Psychology*, *96*, 778–784.

Vaughan, E. (1977). Misconceptions about psychology among introductory psychology students. *Teaching of Psychology*, *4*, 138–141.

Winer, G. A., Cottrell, J. E., Gregg, V., Fournier, J. S., & Bica, L. A. (2002). Fundamentally misunderstanding visual perception: Adults' belief in visual emissions. *American Psychologist*, *57*, 417–424.