
Does Humor in Course Tests Reduce Anxiety and Improve Performance?

Ronald A. Berk

What was the original name of the word processing software WordPerfect?

- A. WordVeryGood
- B. WordAverage
- C. BadWord
- D. ReallyBadWord
- E. WurdWitoutSpelchek

Should this type of item be included on course tests? Does the empirical research on humor indicate that humorous items decrease anxiety and increase test performance? How do you even use humor appropriately in tests? The purpose of this article is to provide an up-to-date summary of the research related to humor in testing, to present new evidence on the effectiveness of humor in constructed-response items, and to describe specific strategies for using humor in tests.

Two research domains exist in which evidence has been collected to determine the value of humor in course tests: first, the psychological and physiological effects of humor, and second, the effects of humor in testing. Brief reviews of these domains follow.

Psychophysiological Evidence

Among the numerous psychological/physiological effects of humor that I have

Ronald A. Berk is a professor of biostatistics and measurement and assistant dean for teaching at the School of Nursing, at Johns Hopkins University, Baltimore, Maryland.

reviewed (Berk 1997, 1998c, 1999, in press), the most pertinent to the conditions of testing in a college classroom is the reduction of anxiety, tension, and stress. There is probably no other time throughout a semester when those negative emotions are at their peak in students' minds and bodies as in the first half hour of an exam.

Humor's primary psychological function is detachment, but it is used as an adaptive coping mechanism (Kline 1907). Using humor in a testing situation produces a cognitive shift in perspective (Dixon 1980; Kuhlman 1984; O'Connell 1976) that allows students to distance themselves from the immediate threat—the test. Humor can reduce the negative feelings that would normally occur, namely, anxiety, tension, and stress. It promotes an objectivity that buffers the negative response (Galloway and Cropley 1999) and provides a sense of empowerment over the testing situation (Sherman 1998; Thorson 1985; Wooten 1996) that can improve students' mental functioning and performance.

The research evidence on the psychological benefits concentrates on three outcomes: first, anxiety reduction (Cann, Holt, and Calhoun 1999; DeSpelder and Strickland 1983; Doris and Fierman 1956; Dworkin and Efran 1967; Liechty 1987; Nezu, Nezu, and Blissett 1988; Singer 1968; Yovetich, Dale, and Hudak 1990); and second, tension reduction (Brill 1940; Goodheart 1994; O'Connell 1960; Wooten 1996). The third is stress

reduction (Bizi, Keinan, and Beit-Hallahmi 1988; Dixon 1976, 1980; Labott and Martin 1987; Lefcourt et al. 1995; Lefcourt and Martin 1986; Lefcourt and Thomas 1998; Martin et al. 1993; Martin and Lefcourt 1983; Nezu et al. 1988; Prerost 1988; White and Camarena 1989; Wooten 1996). When humor permits students to deal more effectively with an adverse experience such as taking a test, the physical response is usually laughter in some form. It may be a smirk, smile, giggle, chortle, or cackle, but rarely convulsive laughter (Kuhn 1994). The laughter translates into physiological effects throughout the entire body (Fry 1992).

The psychological effects reduce physical stress. Humor produces an emotional response that decreases stress, and laughter stimulates a physiological effect that decreases stress hormones such as serum cortisol, dopac, and epinephrine, as well as growth hormone levels in the blood (Berk et al. 1988; Berk, Tan, Fry et al. 1989; Berk, Tan, Napier, and Eby 1989; Fry 1971, 1984, 1992). These effects collectively furnish the physiological evidence for the stress reduction benefit of laughter.

Evidence on Humor in Testing

The empirical research on the effects of humor in testing has been critically reviewed by McMorris, Boothroyd, and Pietrangelo (1997). They found only nine investigations of humor in college testing. All were conducted with students in

undergraduate psychology classes. Seven used humor in multiple-choice items (one of those also included short-answer items); one used anagrams with cartoons, and the last incorporated humor in written dialogue between therapist and client.

McMorris, Boothroyd, and Pietrangelo (1997) concluded that these studies provide insufficient evidence for using humor in exams as a method for reducing anxiety and stress and improving performance, despite students' self-reported preferences for humor. Only the research by Smith et al. (1971) and Hedl, Hedl, and Weaver (1981) reported positive effects of humor on anxiety and stress reduction, respectively. Considering the limitations of many of the investigations and the complexity of measuring interactions between humor in tests and other variables, McMorris, Boothroyd, and Pietrangelo (1997) rendered the following verdict:

Our own personal view at this juncture is to encourage the use of humor in tests, especially if instruction has included use of humor, the test has either no time limit or a very generous one, the humor is positive and constructive, the humor is appropriate for the group, test takers come from the same culture as the item writer, and the test developer feels comfortable in using humor. (295)

Effectiveness of Humor

Given the paucity of evidence on the effects of humor in tests as well as the limited focus and conditions of previous studies (i.e., primarily multiple-choice items with content-irrelevant humor in undergraduate psychology classes), I designed a different study of humor to target (a) constructed-response item formats, but also include multiple-choice and matching formats, (b) content-relevant humor that is an integral part of what an item is measuring, (c) undergraduate and graduate courses in statistics, (d) day and evening graduate courses, (e) students' perceptions of the extent to which humor is effective in reducing their test anxiety and improving their performance, and (f) the reliability and validity of these perceptions with multiple cross-validation samples of different students over six years.

Method

Seventeen convenience samples of students ($N = 695$) enrolled in six under-

graduate and eleven graduate introductory (level one) statistics courses at the Johns Hopkins University's (JHU) School of Nursing were selected for this study. These two- and three-credit courses met for fifteen weeks during each regular semester (not summer) from fall 1994 through fall 1999. The accreditation board requires all students to satisfy these course requirements in the baccalaureate and master's degree programs.

The class sizes ranged from twenty-one to one hundred students. The typical undergraduate in the statistics courses is twenty-six to twenty-seven years old, in his/her junior year of the program, Caucasian (72–86 percent), and female (87–92 percent). African Americans make up 5–10 percent of the students in these courses, Latinos, 2.5–5.0 percent, and Asians 2.5–8.5 percent. Only a small percentage of the undergraduates (24 percent) have had any previous statistics courses that might be used to waive the statistics requirement. Those enrolled in the JHU course would not have had a prior statistics course.

By comparison, the graduate statistics courses comprise advanced undergraduates (up to 60 percent in some classes), master's-level students (30–60 percent), and special students (up to 35 percent) who have not been admitted to any program. These graduate students are older ($M = 32.89$ years old) and more diverse ethnically (64–87 percent Caucasian, 9–16 percent African American, and 3–14 percent Asian) but are still predominantly female (81–91 percent). Further, usually a majority of the graduate students have taken at least one undergraduate statistics course (43–64 percent), and a few (10–30 percent) have had two or three undergraduate and graduate statistics courses.

There are two statistics exams in each course that will have been completed at the point when the students evaluate the humor used in the course. The exams cover descriptive statistics as well as correlation and regression. (The final exam, which is given after the humor evaluation, measures experimental design statistics, such as power analysis, t -tests, analysis of variance, and chi-square.) All exams were open book/open notes with calculators required. No memorization was necessary, and no knowledge-level

items were included. The tests measured students' abilities to understand, apply, and analyze statistical methods within a research context. The items were intended to simulate the actual decisions a researcher would make.

Humor was put into the content of three different item formats: context-dependent multiple-choice, matching, and constructed-response problem-solving. Although the largest percentage of items on all exams were constructed-response (62–90 percent) for both undergraduate and graduate courses, the other formats were also represented as appropriate: multiple-choice (5–10 percent) and matching (15–32 percent). Examples of some of these items are presented in the appendix to this article on page 156.

I administered the Humor Effectiveness Evaluation, developed six years ago for a study of humor strategies (Berk 1996; Berk and Nanda 1998), to students at the end of the last class (before the final exam). They completed it anonymously, and I collected and hermetically sealed the evaluations to preserve confidentiality until final grades were submitted to the registrar's office. The same questionnaire and procedures were used for both undergraduate and graduate courses.

Only one item on two of the three subscales focused on "humorous material on exams." The students were asked to rate the effectiveness of the strategy in producing two outcomes, "to reduce your anxiety on the exam" and "to perform your best on the exam." The students' perceptions of effectiveness were graduated on a scale of *ineffective*, *somewhat effective*, *moderately effective*, *very effective*, and *extremely effective*. These were holistic ratings of effectiveness. Students were not requested to rate the effectiveness of the humor in the different item formats. Their ratings reflected the overall impact of the humor on their anxiety and performance.

The students' ratings of effectiveness for each class were analyzed by computing median ratings for the *anxiety* and *performance* outcomes or criteria. The choice of the median as the appropriate statistic (as opposed to the mean) was based on the ordinality of the qualitative effectiveness anchors and the negatively skewed distributions of the item scores.

The use of means, which assumes an interval scale, would have produced biased results with these distributions.

Results

The median student ratings of "humorous material on exams" for the six undergraduate and eleven graduate statistics classes are shown in table 1. The class sizes reported are the response rates (61–97 percent) of the classes in which the questionnaires were administered and collected. The median ratings are expressed in terms of the verbal scale for ease of interpretation rather than their quantitative code. The ratings for all seventeen classes ranged from *moderately effective* to *extremely effective* for both criteria. In other words, the ratings collectively indicate that the students typically felt that the humor was effective in reducing their anxiety and helping them to perform their best on the exams.

The highest effectiveness rating, *extremely effective*, was found for the majority of undergraduate and graduate classes for "reducing test anxiety." This humor effect was more pronounced than the effect on "performing your best."

Those findings are consistent with the voluminous psychological research evidence reviewed previously on the anxiety reduction effects of humor as well as the results of the Smith et al. (1971) and Hedl, Hedl, and Weaver (1981) studies.

The lowest effectiveness rating by any class was *moderately effective* for "performing your best" by the fall 1999 evening graduate class ($n = 33$). Despite this single *moderately effective* median rating, there were no discernible differences between day and evening graduate class ratings over the six years.

The analysis of class ratings of the undergraduate and graduate classes indicated a small (5–15 percent) but consistent percentage of students in the graduate classes who rated the use of humor on the tests for both anxiety reduction and improved performance as *ineffective*. That pattern was not evident in the undergraduate classes.

To probe possible explanations for these trends, additional items were added to the rating scale administered in the fall 1999 courses: (a) The humorous items were distracting, or (b) the humorous items were confusing. Using a six-point

strongly agree to *strongly disagree* scale, the graduate class reported 23 percent agreeing with the first statement and 19 percent agreeing with the second. By comparison, the undergraduate class contained only 2 percent (one student) who agreed that humor was distracting and none who said that it was confusing. Consistent with these undergraduate-graduate discrepancies in ratings were the students' responses to the statement that humorous items should be included on the next test. The same 19 percent of the graduate and 2 percent of the undergraduate students disagreed with that suggestion.

The preliminary evidence on the fall 1999 classes suggests that a few graduate students may be bothered by humorous test items and perceive them as ineffective in reducing anxiety and improving their performance. Perhaps graduate students' more serious approach to test taking cannot be positively affected by humor.

Discussion

The evidence presented in table 1 is compelling for a number of reasons. First, it indicates that most students feel that humor makes a difference in their test performance; second, this effect was demonstrated with three item formats measuring "real-life" higher-order thinking skills; third, the effect is not idiosyncratic to just one class or level of students; fourth, the effect is consistent over time (six years) and accurate for different cross-validation samples; and finally, the effect is general across undergraduate courses, day and evening graduate courses, and a range of class sizes, students, and exams.

Although these inferences are restricted to the specific statistics courses, the humor strategies tested are not peculiar to statistics. Unfortunately, the holistic ratings of the humor effectiveness survey did not partition the possible differential effects of constructed-response, multiple-choice, and matching formats. Nor was it possible to isolate the impact of other forms of humor, such as humor in the title, directions, and notes on the last page of the test. Those limitations of the dependent variable need to be addressed in future studies. However, it is possible that students may not be able to recall those specifics or be capable of discrimi-

Table 1.—Median Student Ratings of Humorous Material on Exams for Six Undergraduate and Eleven Graduate Statistics Classes (1994–1999)

Class (size)	Criteria	
	To reduce your anxiety on the exam	To perform your best on the exam
Undergraduate		
1994 ($n = 30$)	Very effective	Very effective
1995 ($n = 42$)	Extremely effective	Very/ext. effective
1996 ($n = 29$)	Extremely effective	Extremely effective
1997 ($n = 39$)	Extremely effective	Very effective
1998 ($n = 32$)	Extremely effective	Very effective
1999 ($n = 51$)	Extremely effective	Very effective
Graduate		
1994 ($n = 44$)	Very effective	Very effective
1995 ($n = 68$)	Extremely effective	Very effective
1995 ($n = 21$)	Very effective	Very effective
1996 ($n = 46$)	Extremely effective	Very effective
1996 ($n = 36$)	Extremely effective	Extremely effective
1997 ($n = 58$)	Extremely effective	Very effective
1997 ($n = 15$)	Extremely effective	Extremely effective
1998 ($n = 56$)	Very effective	Very effective
1998 ($n = 27$)	Extremely effective	Very effective
1999 ($n = 68$)	Extremely effective	Very effective
1999 ($n = 33$)	Very effective	Moderately effective

nating accurately among different item formats by the end of the semester.

Strategies for Using Humor in Tests

The research described in the following sections defines the types and forms of humor investigated, some with illustrative items. Unfortunately for professors who wish to include humor in their tests, there is no article, chapter, or comprehensive book that provides step-by-step guidelines on how to use humor appropriately. In many cases, professors have incorporated humor inappropriately by default, which unfortunately has resulted in a few college departments' banning the use of humor in tests.

Although inappropriate use may not be intentional, it can have a negative impact on the students and on the psychometric properties of the test. Instead of obtaining the desired positive effects of reducing anxiety and stress and improving performance, inappropriate use can be offensive, distracting, and anxiety producing, which can decrease test performance. It can also adversely affect the validity of the items and the reliability of the scores.

In an effort to maximize the potential benefits of using humor, I present in this section a variety of strategies I have tested over the past thirteen years in my classes and in my humor research on "effectiveness" within the past six years. More examples of these techniques are given in the appendix to this article.

There are at least four major techniques, only one of which deals with the actual test items. They include the following: incongruous descriptors under the title of the test, jocular inserts in the directions, humorous notes on the last page, and humor in the test items. The first three are described and illustrated elsewhere (Berk 1998c). I will examine the eight strategies for infusing humor into test items, which can be separated into two categories, content-irrelevant and content-relevant. The first set of techniques is the simplest to develop and apply because the humor is added to the existing test; the second involves incorporating humor into the content of the items.

Content-Irrelevant Strategies

These strategies are most appropriate for multiple-choice format items. Step

one is to construct a completely serious, boring test. Suppose it consists of fifty four-choice items. The humor that can be inserted is an add-on to the "nonhumorous" test. That is, the humor will not affect the structural integrity, validity, or reliability of any item or of the total score. Regardless of what humor is used, the original fifty-item test is preserved, and the outcomes being measured are not affected.

Add humorous distracters to several items. If there are four content choices for each item, tack on a humorous choice "E" that is irrelevant. That choice should be so ridiculous and outrageous that no student could possibly consider it as the correct answer. It should be the only implausible distracter in the item. An example of this technique is illustrated in the appendix on page 156.

The humorous fifth choice should be added to several items throughout the test to help reduce anxiety and motivate students to "plow through" all the items. One advantage of this strategy is that the added distracters require no major increase in reading time. In other words, the added fifth choice can be used in both speed and power tests. If students are pressed for time to finish the test, the humor in the distracters should not interfere.

A preview in the test directions announcing the appearance of these items is advisable to convey that there is something positive about the testing experience and to warn those few super-serious students with the sense of humor of wood varnish that choice "E" is just a joke and not to be chosen as a correct answer. The statement may take the following form:

This test consists of fifty multiple-choice items. Most of the items have four choices (A, B, C, or D); however, there are several items sprinkled throughout the test with a humorous choice "E," which is intended solely for your entertainment. Any other use, such as for the correct answer, is strictly prohibited without prior written consent.

Add humorous items throughout the test. Instead of tinkering with the internal structure of the items by adding humorous distracters, you can just add a few humorous items to all of the serious ones. Because this strategy requires additional

reading time, it should only be considered for tests in which students have adequate time to complete all of the items. If speed is important and some students are rushed to finish the test, no other items should be added.

Humorous items function most effectively in power tests in which what a student knows or can demonstrate is more critical than how fast he or she can answer. Those tests are often called criterion-referenced, mastery, or competency-based. Suppose five humorous items were inserted in a fifty-item test. Each one might appear after every ten serious items. These five would not be counted in the total score so that the validity and reliability of the original test scores remain intact. The humorous items may be written on any topic irrelevant to the specific content of the test. In fact, they may be opinion items or other items with no "real" correct answer.

The reasons for scattering the humorous items throughout the test are the same as for sprinkling the items with humorous distracters. It is also recommended that a statement similar to the preceding one appear in the directions, for example:

There are several humorous items sprinkled throughout the test, which are intended solely for your entertainment. Any other use, such as for food, shelter, or wedding invitations, is strictly prohibited. Because these items will not be counted in your final score, it is your decision whether to answer them or not.

Add humorous distracters and items. One can use a combination of the two preceding strategies. A variety of humorous multiple-choice item examples can be found in Berk (1998a, 1998c). However, do not overdo it. Too much humor can be distracting. Power tests afford greater latitude in the use of humor than speeded tests.

Content-Relevant Strategies

These strategies involve injecting humor into the actual item content; thus it is not an add-on to be dismissed by the student. This form of humor can appear in constructed-response as well as multiple choice and matching items.

Insert humor into the stem of a multiple-choice item. If knowledge or application of a concept is being tested, the

example in the stem may be humorous while all of the content choices are serious. Consider the following item:

The Artist formerly known as MS-DOS is:

- A. Linux
- B. X Window
- *C. Windows
- D. Unix
- E. Mac OS

Insert humor into the choices of a multiple-choice item. The previous strategy can be reversed by placing serious content in the stem and writing humorous, but plausible, choices:

What is the most appropriate politically correct term for *ex-spouse* (not necessarily yours)?

- A. cerebrally challenged
- B. parasitically oppressed
- *C. insignificant other
- D. socially misaligned

Insert humor into the stems of a matching item set. Because a matching set is just a collection of multiple-choice items put together, the stems in Column 1 may be humorous, while the Column 2 choices are serious. This strategy can be used with the two popular types of matching: one with four or five choices that can be selected more than once, or one with a long list of choices (three or four more than the number of stems) each of which can be selected only once. The use of humor in these two types of matching is illustrated in the appendix.

Insert humor into context-dependent material. When a student is presented with verbal or visual/pictorial stimuli to interpret before answering a set of multiple-choice items (a.k.a. testlet), that material provides an opportunity for humor. The verbal material may be a hypothetical humorous observation in a problem, scenario, vignette, or case study. The visual/pictorial material can take the form of a cartoon; humorous picture or photograph; or humorous variables or data in a table, chart or graph, map, figure, or diagram. These forms of humor require more imagination than any of the preceding, but they can be just as effective as nonhumorous material in

measuring higher-order thinking skills, such as understanding, critical thinking, reasoning, problem solving, decision making, daydreaming, and country line dancing. One example is shown below:

A group of patients with flu-like symptoms was randomly split into two samples. Sample one received lozenge megadoses of *Cold-Wheeze* and the other received comparable doses of *Zincy-Winky* (one lozenge has so much zinc, you couldn't get through the airport metal detector). After one week, the patients rated the effectiveness of these two medications in reducing their symptoms. These ratings ranged from 0 (I still feel like a zombie) to 25 (I'm flying high like Rocky!). Determine whether CW was superior to ZW.

1. What is the most appropriate research hypothesis?

A. *Cold-Wheeze* patients will rate their medication significantly more effective than *Zincy-Winky* patients.

B. *Zincy-Winky* patients will rate their flu-lozenge more effective than *Cold-Wheeze* patients.

C. There will be a significant difference in the effectiveness of *Cold-Wheeze* and *Zincy-Winky* in reducing flu-like symptoms.

D. There will be no difference in *Cold-Wheeze* and *Zincy-Winky's* effectiveness in reducing flu-like symptoms.

2. What statistical test should be used?

- A. *t* test (pooled variance)
- B. *t* test (separate variance)
- C. *t* test (repeated measures)
- D. Chi-square test of independence

Insert humor into constructed-response stimuli. The humor strategy for constructed-response format items is virtually identical to the preceding strategy. The difference between the items is in the form of the student's response. Instead of providing answers to a testlet, the student constructs his or her own answer to an open-ended question(s) about the material presented. An example of this type of question is provided in the appendix.

Evaluate the Effectiveness of Your Humor

Whether the humor on the test is effective can be measured by asking your students. Solicit their reactions to the humor after the first test, so that there is time to make adjustments on subsequent tests.

The easiest method is to distribute a three-by-five-inch index card to each student and ask them to answer any of the following yes-no questions: Did you like the humorous items on the test? Were they too distracting? Did they reduce your test anxiety? Did they help you perform your best on the test? Should humorous items be included on the next test?

Also ask the students for comments or humor suggestions. Make certain no names are written on the cards. Anonymity is essential to obtain truthful responses. And, yes, you will get feedback, probably more than you ever expected. Their responses may run the gamut from "Knock off the humor" to "Keep it up; the more the better." The majority or median opinion should guide your use of humor in the next test. Keep in mind that your credibility is at stake. Tell the students the results of their index-card evaluation and your justification for using or not using humor on subsequent course tests.

Conclusion

After reviewing the available research and strategies for injecting humor into test items, how can the question posed in the title of this article be answered? One way is to simply say, "Yes. Humor *does* reduce anxiety and improve performance." But that will not do because there are too many qualifiers.

This article provides six conclusions, therefore, about the use of humor in course tests: (1) there is psychological and physiological research evidence that humor reduces anxiety, tension, and stress; (2) no one has documented the frequency, duration, and level of laughter (e.g., giggle, chortle, or convulsive hysteria) that occurs during testing, although students do agree that some tests are more humorous than others; (3) the few studies conducted on the use of humor in testing yield insufficient and inconsistent results; (4) the findings of the six-year study reported in this article indicate that undergraduate and graduate students feel humor is effective in reducing their anxiety and making it possible to perform their best on exams; (5) there are four major techniques for using humor in tests, only one of which involves the items, and (6) there are at least three content-irrelevant and five

content-relevant strategies that can be used to infuse humor into traditional multiple-choice, context-dependent multiple-choice, matching, and constructed-response items.

The research based on student perceptions strongly suggests that humor is

effective. But when appropriate measures of anxiety, stress, and achievement are used as dependent variables, the effects are not so definitive. There is a glaring need for well-designed studies of the effectiveness of humor in a variety of item and test formats. Perhaps one deter-

rent to previous efforts has been the lack of sources that describe how to use humor in tests appropriately. I hope that the didactic section of this article will correct that deficiency and facilitate both future applications of humor and the research on their effectiveness.

APPENDIX

These are a few examples of the content-irrelevant and content-relevant strategies that I have tested in my classrooms over the past thirteen years and in my research on the effectiveness of humor.

One method of injecting humor into an exam is to add humorous distractors to several items. An example of this technique would be:

Who wrote "The Love Song of J. Alfred Prufrock"?

- A. T. S. Eliot
- B. C. S. Lewis
- C. A. E. Housman
- D. J. D. Salinger
- E. L. L. Bean

Although "L. L. Bean" follows the same content format as the other choices, there are generic choices that can be added to any item, that are sure to bring at least a smile or a pleasant diversion.

Rather than altering the internal structure of test questions or answers, a second technique would be to sprinkle entirely humorous items throughout the exam. One such item is given below:

Which "oxymoron" actually makes sense?

- A. Pretty ugly
- B. Resident alien
- C. Act naturally
- D. Passive aggressive
- E. Soft rock

It has also been argued that "humor in academia" is an oxymoron with three words (Berk 1998b).

The following two examples are content-relevant strategies that insert humor into the stems of matching item sets. Type 1 represents a test item with many possible answers that can be used repetitively, whereas Type 2 allows the responder to match each item only once.

Type 1

Directions: Match the variables or questionnaire items in Column I with their HIGHEST level of measurement in Column II. Mark your answers (A, B, C, or D) in spaces 1-10 of your answer sheet. Each response in Column II may be used once or more than once.

Column I

Column II

- | | |
|---|--|
| <p>___ 1. Wait time to see your doctor</p> <p>10 minutes or less</p> <p>More than 10 but less than 30 minutes</p> <p>Between 30 minutes and an hour</p> <p>More than an hour and less than a day</p> <p>I'm still waiting</p> <p>___ 2. Years of experience as a stand-up comic</p> <p>___ 3. Native language</p> <p>English</p> <p>Spanish</p> <p>Latin</p> <p>Greek</p> <p>Psychobabble</p> <p>Other</p> <p>___ 4. Degree of frustration</p> <p>Totally give up</p> <p>Might give up</p> <p>Thinking about giving up</p> <p>Refuse to give up</p> <p>Don't know meaning of "give-up"</p> <p>___ 5. Scores on the "Where's Waldo" final exam (0-100)</p> <p>___ 6. Symptoms of exposure to statistics</p> <p>Vertigo</p> <p>Nausea</p> <p>Vomiting</p> <p>Numbness</p> <p>Hair Loss</p> <p>___ 7. Length of your big toe</p> <p>___ 8. Ratings on the Lack of Quality Inn Scale (0-60)</p> <p>___ 9. Health status leaving Johns Hopkins Hospital (dead or alive)</p> <p>___ 10. Quantity of blood consumed by Dracula per night</p> | <p>A. nominal</p> <p>B. ordinal</p> <p>C. interval</p> <p>D. ratio</p> |
|---|--|

Type 2

Directions: Match the famous classical music compositions in Column I with the guys who probably wrote that

APPENDIX

kind of stuff in Column II. Mark your answers (A–M) in spaces 1–10 of your answer sheet. Each choice in Column II may be used once or not at all.

<u>Column I</u>	<u>Column II</u>
___ 1. Les Quaaludes	A. Offenbach
___ 2. Pictures of an Exhibitionist	B. Rimsky-Korsakov
___ 3. Marche Slob	C. Chopin
___ 4. Orpheus in His Underwear	D. Orff
___ 5. Le Coq Au Vin	E. Liszt
___ 6. Pathetic	F. Stravinsky
___ 7. The Firefly	G. Debussy
___ 8. Peter and the Moose	H. Prokofiev
___ 9. Capriccio Escargot	I. Mussorgsky
___ 10. A Midsummer Night on Elm Street	J. Beethoven
	K. Tchaikovsky
	L. Rachmaninoff
	M. Mendelssohn

The final example represents a way to inject humor into constructed-response stimuli. Students must construct their own answer to the open-ended question. This item is a statistics simulation problem unlike those found in most textbooks:

One group of twenty-five students volunteered to participate in a study about the quality of TV programs. Each student was asked, "Yo, student! Have you ever watched *The*

Practice, *Chicago Hope*, or *Ally McBeal*?" The typical response was: "Are you nuts? We're students. Heellooo! We don't have a life!" Anyway, eighteen students had never seen any of those programs. Those eighteen were randomly divided into three samples of six students each and then randomly assigned to watch three episodes of one of those programs. Each student then rated the overall writing quality of the episodes according to level of "creativity" on a 15-point scale (0 = *trite, dull, insulting*, to 15 = *imaginative, cutting-edge creativity*). The ratings by group are shown below. Determine which show should receive an Emmy for creative writing.

<u>The Practice</u>	<u>Chicago Hope</u>	<u>Ally McBeal</u>
2	1	11
4	3	12
6	5	13
8	7	14
10	9	14
12	11	15

- a. Research hypothesis:
- b. Null hypothesis:
- c. Statistical test:
- d. Significance level:
- e. Sampling distribution:
- f. Computation of statistic:
- g. Conclusion:

REFERENCES

- Berk, L. S., S. A. Tan, S. Nehlsen-Cannarella, B. J. Napier, J. E. Lewis, J. W. Lee, and W. C. Eby. 1988. Humor associated laughter decreases cortisol and increases spontaneous lymphocyte blastogenesis. *Clinical Research* 36:435A.
- Berk, L. S., S. A. Tan, W. F. Fry, Jr., B. J. Napier, J. W. Lee, R. W. Hubbard, and J. E. Lewis. 1989. Neuroendocrine and stress hormone changes during mirthful laughter. *American Journal of the Medical Sciences* 298:390–6.
- Berk, L. S., S. A. Tan, B. J. Napier, and W. C. Eby. 1989. Eustress of mirthful laughter modifies natural killer cell activity. *Clinical Research* 37:115A.
- Berk, R. A. 1996. Student ratings of 10 strategies for using humor in college teaching. *Journal on Excellence in College Teaching* 7(3): 71–92.
- . 1997. Top ten strategies for using humor as an effective teaching tool. In *Selected papers and awards from the Eighth National Conference on College Teaching and Learning*, ed. J. A. Chambers, 9–27. Jacksonville, Fla.: Florida Community College and Jacksonville Foundation.
- . 1998a. A humorous account of 10 multiple-choice item flaws that clue test-wise students. *Journal on Excellence in College Teaching* 9(2): 93–117.
- . 1998b. Is humor in academia an oxymoron? *Quality in Higher Education* 7(1): 1–2.
- . 1998c. *Professors are from Mars, students are from Snickers: How to write and deliver humor in the classroom and in professional presentations*. Madison, Wisc.: Magna Publications.
- . 1999. Laughter therapy. *The Johns Hopkins Health Insider* 2(4): 1, 8.
- . In press. The active ingredients in humor: Psychophysiological benefits/risks for older adults. *Educational Gerontology: An International Journal*.
- Berk, R. A., and J. P. Nanda. 1998. Effects of jocular instructional methods on attitudes, anxiety, and achievement in statistics courses. *HUMOR: International Journal of Humor Research* 11:383–409.
- Bizi, S., G. Keinan, and B. Beit-Hallahmi. 1988. Humor and coping with stress: A test under real-life conditions. *Personality and Individual Differences* 9:951–6.
- Brill, A. A. 1940. The mechanism of wit and humor in normal and psychopathic states. *Psychiatric Quarterly* 14:731–49.
- Cann, A., K. Holt, and L. G. Calhoun. 1999. The roles of humor and sense of humor in responses to stressors. *HUMOR: International Journal of Humor Research* 12(2): 177–93.
- DeSpelder, L., and A. Strickland. 1983. The last dance: Encountering death and dying. *Death Studies* 9:201–16.
- Dixon, N. F. 1976. *On the psychology of military incompetence*. London: Cape.
- . 1980. Humor: A cognitive alternative to stress? In *Stress and anxiety* (Vol. 7), ed. I. G. Sarason and C. D. Spelberger, 281–9. Washington, D.C.: Hemisphere.
- Doris, J., and E. Fierman. 1956. Humor and anxiety. *Journal of Abnormal and Social Psychology* 53:59–62.
- Dworkin, E. S., and J. S. Efran. 1967. The angered: Their susceptibility to varieties of humor. *Journal of Personality and Social Psychology* 6:233–6.
- Fry, W. F., Jr. 1971. Mirth and oxygen saturation levels of peripheral blood. *Psychotherapy and Psychosomatics* 19(1): 76–84.
- . 1984. Learning with humor. Paper presented at the annual meeting of the International Conference on Humor, Tel Aviv, Israel.

- . 1992. The physiological effects of humor, mirth, and laughter. *Journal of the American Medical Association* 267(4): 1857–8.
- Galloway, G., and A. Cropley. 1999. Benefits of humor for mental health: Empirical findings and directions for further research. *HUMOR: International Journal of Humor Research* 12(3): 301–14.
- Goodheart, A. 1994. *Laughter therapy*. Santa Barbara, Calif.: Stress Less Press.
- Hedl, J. J., Jr., J. L. Hedl, and D. B. Weaver. 1981. The effects of humor on anxiety and performance. Paper presented at the annual meeting of the American Educational Research Association, 21–25 April, Los Angeles, Calif.
- Kline, L. W. 1907. The psychology of humor. *American Journal of Psychology* 18:421–41.
- Kuhlman, T. 1984. *Humor and psychotherapy*. Homewood, Ill.: Dorsey Press.
- Kuhn, C. C. 1994. The stages of laughter. *Journal of Nursing Jocularly* 4(2): 34–5.
- Labott, S. M., and R. B. Martin. 1987. The stress-moderating effects of weeping and humor. *Journal of Human Stress* 13:159–64.
- Lefcourt, H. M., and R. A. Martin. 1986. *Humor and life stress*. N.Y.: Springer-Verlag.
- Lefcourt, H. M., K. Davidson-Katz, R. Shepherd, and M. Phillips. 1995. Perspective-taking humor: Accounting for stress moderation. *Journal of Social and Clinical Psychology* 14:373–91.
- Lefcourt, H. M., and S. Thomas. 1998. Humor and stress revisited. In *The sense of humor: Explorations of a personality characteristic* (Humor Research Series, Vol. 3), ed. W. Ruch. Berlin, Germany: Morton de Gruyter.
- Liechty, R. D. 1987. Humor and the surgeon. *Archives of Surgery* 122:519–22.
- McMorris, R. F., R. A. Boothroyd, and D. J. Pietrangelo. 1997. Humor in educational testing: A review and discussion. *Applied Measurement in Education* 10(3): 269–97.
- Martin, R. A., and J. P. Dobbin. 1988. Sense of humor, hassles, and immunoglobulin A: Evidence for a stress-moderating effect of humor. *International Journal of Psychiatry in Medicine* 18:93–105.
- Martin, R. A., N. A. Kuiper, L. J. Olinger, and D. A. Dance. 1993. Humor, coping with stress, self-concept, and psychological well-being. *HUMOR: International Journal of Humor Research* 6(1): 89–104.
- Martin, R. A., and H. M. Lefcourt. 1983. The sense of humor as a moderator of the relationship between stressors and mood. *Journal of Personality and Social Psychology* 45:1313–24.
- Nezu, A. M., C. M. Nezu, and S. E. Blissett. 1988. Sense of humor as a moderator of the relation between stressful events and psychological distress: A prospective analysis. *Journal of Personality and Social Psychology* 54:699–714.
- O'Connell, W. E. 1960. The adaptive function of wit and humor. *Journal of Abnormal and Social Psychology* 61:263–70.
- . 1976. Freudian humour: The Eupsychia of everyday life. In *Humour and laughter: Theory, research and applications*, ed. H. C. Foot and A. J. Chapman, 313–30. London: Wiley.
- Prerost, E. 1988. Use of humor and guided imagery in therapy to alleviate stress. *Journal of Mental Health Counseling* 10(1): 16–22.
- Sherman, K. M. 1998. Healing with humor. *Seminars in Perioperative Nursing* 7(2): 128–37.
- Singer, D. L. 1968. Aggression arousal, hostile humor, and catharsis. *Journal of Personality and Social Psychology Monograph Supplement* 8:1–14.
- Smith, R. E., J. C. Ascough, R. F. Ettinger, and D. A. Nelson. 1971. Humor, anxiety, and task performance. *Journal of Personality and Social Psychology* 19:243–6.
- Thorson, J. A. 1985. A funny thing happened on the way to the morgue: Some thoughts on humor and death, and a taxonomy of the humor associated with death. *Death Studies* 9:201–16.
- White, S., and P. Camarena. 1989. Laughter as a stress reducer in small groups. *HUMOR: International Journal of Humor Research* 2(1): 73–9.
- Wooten, P. 1996. Humor: An antidote for stress. *Holistic Nursing Practice* 10(2): 49–56.
- Yovetich, N. A., J. A. Dale, and M. A. Hudak. 1990. Benefits of humor in reduction of threat-induced anxiety. *Psychological Reports* 66:51–8.