Intuition: BYU Undergraduate Journal of Psychology, is published annually during the Fall Semester by the Brigham Young University Psychology Department.

Intuition is managed, edited, and designed by undergraduate students enrolled at Brigham Young University. Individuals who are interested in joining the staff are encouraged to contact the editor-in-chief at byupsychjournal@gmail.com. Class credit is available for staff members. Individuals interested in submitting articles to be considered for publication should read the submission guidelines located on the last page of the current issue.

Articles and other content in this issue do not necessarily represent the views or opinions of Brigham Young University, The Church of Jesus Christ of Latter-day Saints, or Intuition editors. Any comments should be directed to the Editor-in-Chief, 1001 SWKT, BYU, Provo, UT, 84602, or e-mail: byupsychjournal@gmail.com.

Intuition would especially like to thank Harold Miller for his guidance in developing this journal. The efforts of Mel Thorne, Linda Hunter Adams, and the Psychology Department faculty members who generously provided reviews are likewise appreciated.

Additional information can be found on our Web site at http://intuition.byu.edu/index.html.
# Contents

From the Editor

Michael John Redd

A Brief History of Psychology at Brigham Young University (Part 2, 1976–2006)

Brandon L. Roberg

As Part 2 of a series on the history of psychology at Brigham Young University, this article covers the years 1976–2006. Part 1 covers the years 1876–1975 and can be found in Volume 1 of Intuition.

An Interview with Dr. Allen Bergin

Brandon L. Roberg

Dr. Allen Bergin, a BYU emeritus professor, was one of the most eminent members of the BYU Department of Psychology. His influence can be seen throughout the discipline of psychology. This interview took place via e-mail during June 2006.

Publishing in Psychology: An Overview for Undergraduate Students

Dennis Wendt Jr.

This article relies on the expertise of former and current journal editors to introduce undergraduate students to important issues related to publishing in academic psychology journals. Includes a list of publication helps for undergraduate psychology students.

Tokenism in Non-Native English Speakers: Does It Affect Their English Testing Abilities?

David Shwalb, Callie Beck, Heidi Richards, Rachel Fichtner, Ryan Alexander

This study was conducted with 16 international non-native English speaking undergraduate students to examine the impact of tokenism on such students' test-taking performance.

The Effects of Semantic Relationships on the Irrelevant Sound Effect

Brady Firth, Abigail Clark, Shawn Lelle, Jesse Anderson

This study investigated whether the degree of semantic relationship between audible distracter words and visually displayed to-be-remembered words impacts short-term memory recall.

The Influence of Perceived Parenting Styles on the Degree of Adult Children's Allocation of Punishment

Joseph Gillis, Julie Berry, Jacob Douglas, Erin Evans

The degree and type of punishment that 84 adult participants would allocate for both civic and household offenses was measured. The adult children of authoritative fathers issued significantly more severe punishments with both civic and household offenses.

Retrograde Amnestic Symptoms from Witnessing a Traumatic Event

Marko A. Moreno, Cassie L. Antley, Torrie L. Casebolt, Chesley A. Ward

This study examined whether witnessing a mildly traumatic event may result in retrograde amnestic symptoms. Two groups of participants were surveyed after a brief presentation, one of which witnessed a mildly traumatic event during the presentation.

Evidence-Based Practice Movements in Psychology: Empirically Supported Treatments, Common Factors, and Objective Methodological Pluralism

Dennis Wendt Jr.

Through analysis of three recent movements towards evidence-based practice, the author concludes that EST and common factors movements are both based on "limited, biased conceptions of evidence that constitute a preinvestigatory bias against alternative conceptions and the methods they imply." Objective methodological pluralism is examined as an alternative conception to these two movements.
From the Editor

Michael John Redd

Welcome to the second volume of *Intuition: BYU Undergraduate Journal of Psychology*. We are proud to present this journal and are excited about the content of this latest edition of the journal. In this volume, you will find a new feature that we hope to continue—an interview with a BYU psychology department faculty member. In this case you will get to meet Dr. Allen Bergin, now an emeritus faculty member.

In this volume, you will also see a brief timeline of the BYU Department of Psychology, an editorial about undergraduate students and publishing, four research articles, and one theoretical paper on evidence-based practice. We hope that in this volume you will find something of interest. Please remember that, in the end, the content of the journal depends on the hard-working students who submit articles. If you would like to help improve the quality and breadth of our journal please submit original articles on issues that you think need to be addressed.

Much has happened since the conception of *Intuition: BYU Undergraduate Journal of Psychology* in 2004. The first issue of the journal was published in Fall 2005. Since then, we have been working to advertise the name of *Intuition* and to try to improve the journal. We are also making efforts to increase both the readership of the journal and submissions to the journal.

When I first started as editor-in-chief, my primary focus was on obtaining substantive and meaningful articles, good research, and publishable material. Although that remains my goal for this journal, I have come to appreciate the other opportunities this project has allowed.

The journal has allowed important opportunities to those who worked on the staff and for the authors of the articles submitted. Perhaps the project of publishing *Intuition* will be more beneficial to those of us who have worked on it than those who just read it. Personally, I feel that the project has helped expand my vision, create an interest in me to seek out and learn more about human psychology, and create opportunities to interface with the great staff members, editors, and faculty members that worked on this project.

We feel confident that this student project is important and worthwhile. Our aim throughout has been to help introduce students to the rigorous process of revising and publishing psychological research. We also aim to help all those involved in the project to become better prepared for entrance into graduate school. We hope that our objectives will be better reached as this project continues to mature. We invite all of you to help us with these objectives—to further the study of psychology at BYU—by participating with the journal, submitting your research/theoretical manuscripts, and helping to make this journal a success.
A Brief History of Psychology at Brigham Young University (Part 2, 1976–2006)

As part two of a series on the history of psychology at Brigham Young University, this article covers the years 1976–2006. Part 1 covers the years 1876–1975 and can be found in Volume 1 of Intuition.

1978
Dr. Sally Barlow becomes the first female faculty member of the Department of Psychology and currently continues to teach and research as a professor of psychology and a clinical faculty member.

1980
The Taylor Building serves as a training and research facility, housing several of BYU's clinical training programs, including psychology.

1981
The Department of Psychology moves to the 10th floor of the newly constructed Spencer W. Kimball Tower, where it still resides today.

1990
Dr. Erin D. Bigler returns to BYU, the school where he received both his undergraduate and graduate degrees, and establishes the Brain Imaging & Behavior Lab.

1991
Dr. Michael J. Lambert, a professor of clinical psychology, is asked by Human Affairs International to develop an outcome questionnaire (OQ-45) to improve the quality of mental health treatments. Today the OQ-45 is used on 4 continents and is available in 17 languages.
<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>Dr. Kenneth L. Higbee, professor of psychology, celebrates the 25th anniversary of his book <em>Your Memory, How It Works &amp; How to Improve It</em>. The book is currently published in four languages (Portuguese, Spanish, Mandarin Chinese, and English) and in four countries. It has been continuously in print for 30 years.</td>
</tr>
<tr>
<td>2005</td>
<td>The first annual Fulton Mentored Student Research Conference gives students an opportunity to display the results of their faculty-mentored research projects.</td>
</tr>
<tr>
<td>2005</td>
<td>The first undergraduate journal of the Department of Psychology (<em>Intuition</em>) is published.</td>
</tr>
<tr>
<td>2006</td>
<td>In the winter semester of 2006, there are 987 psychology majors at BYU, making it one of the most popular majors at the university.</td>
</tr>
</tbody>
</table>
How did you decide to pursue a career in psychology?

My early interests were in science and math, which is why I began college at MIT; but being in an intense scientific atmosphere for a year convinced me that I was not cut out for such an occupation full time. This was a surprise and a shock. I then wandered for two subsequent years, taking a wide variety of classes, first at a liberal arts college and then at BYU. During the fall of my junior year, which was at BYU, I was still without a major; but by the end of that quarter I had taken enough psychology in-depth to discover that I loved it. The field brought together in one package the full range of my interests—scientific, philosophical, and humanistic. I gravitated first to experimental psychology and then to the research side of personality and clinical psych, where I found a balanced comfort zone that fit my personality.

How did you decide to enter the master's program at BYU?

After an intense focus on psychology for a year, catching up on the requirements for a major, I decided to apply for doctoral programs in clinical psych, as I realized that my future in psych would be limited without a Ph.D. This plan proved to be unrealistic. I didn't realize how hard it was to get into good clinical programs; I didn't have a solid competitive background in psych, and my grades had suffered during the earlier time of wandering and switching schools. Consequently, I was admitted only to second-tier schools and none of the top ones I preferred. The BYU psych department offered me a generous scholarship, so I decided to stay a year for a master's and develop my qualifications for a first-tier doctoral program. (There were no BYU doctoral programs in 1956.) This proved to be a perfect solution. I finished the master's requirements in one year and I was able to consolidate and expand my knowledge of psychology. I also solidified my understanding of the LDS Church, which I had
What are a few of your memories of working with Carl Rogers?

By the time I finished my Stanford Ph. D. and spent considerable time under the tutelage of Albert Bandura, I realized that the study of personal change was for me. Consequently, the opportunity for a postdoctoral fellowship with Carl Rogers, the father of psychotherapy research, was a dream come true. Carl was all that people imagined him to be: warm, genuine, caring, and gifted in perception and wisdom. He was energetic, hardworking, and creative. He devoted his entire life to the enterprise of facilitating growth in other people via research, theory, practice, teaching, and organizational intervention.

I remember fondly his interest in my wife, Marian, and our children, and the pleasure we had at his home near a lake in Wisconsin, including motorboat rides for the children. I recall vividly our various letters and our personal meetings over the years. On one occasion he gave me a bear hug while congratulating me on my work in spirituality and mental health, an issue over which we had disagreed in earlier years. He said he was pleased that I was speaking straight from “my own gut” and not trying to please anyone else. He always admired self-congruence or honest self-expression regardless of the viewpoint taken; but he had also changed his opinion and by then saw the value of spiritual experience in therapeutic change.

It was also instructive to see this warm, empathic person shift into an authority mode when needed, for instance in disciplining a staff member or rebuking an obnoxious student. There were two dozen grad students and postdocs working under him at the time (1960–61), so he had to be a manager as well as a therapist and an international figure.

What has been one of the highlights of your professional career?

Generally, this would be feeling that I was part of the positive development of psychology and of students, clients, and younger colleagues. Specifically, this past year (2005–2006) has been a highlight as two BYU faculty whom I mentored earlier and supported and collaborated with later became presidents of their respective professional organizations: Michael Lambert as president of the International Society for Psychotherapy Research and Scott Richards as president of the Division of Psychology and Religion in the American Psychological Association. Their achievements are in the two areas I specialized in and for which I received awards from several national organizations. So, it is a double delight to have both my own career and that of my younger esteemed colleagues receive recognition for our work in psychotherapy and in spirituality.

What has been one of the greatest challenges of your professional career?

The greatest challenge was maintaining equanimity and suppressing the impulse to retaliate when I was unfairly criticized or maligned or when false rumors were told about me by colleagues in the profession who strongly disagreed with my views or my actions to promote my position on issues. This included a few colleagues at BYU, which was quite painful to endure. I learned to love my enemies from these experiences and many became treasured friends as a result. I also discovered that I wasn’t always right in my opinions and methods. I also
learned that the Lord loved me and he supported me in many ways, including through his leaders in the Church and at BYU.

What is one psychological topic that you would like to see researched and discussed more thoroughly?

How psychology in all its facets can provide skills, research, spiritually attuned theories, resources, programs, techniques, and findings for advancing the purposes of the LDS Church, which I believe to be one organization in the world that will last indefinitely and has great potential to alter the world for the better, including in collaboration with others of good will.

What has been your greatest conflict between your religious beliefs and mainstream psychology?

Barring for nearly 50 years against the negative influences of naturalism and moral relativism which have pervaded the field and the other behavioral and social sciences in both theory and practice. My writings, research, teaching, and speeches have been a continuing testimony against these views of human nature and in favor of a moral and spiritual perspective.

What advice do you have for up-and-coming psychologists, particularly those of the BYU student body?

Seek first the kingdom of God and not self-aggrandizement or riches, then many blessings will come; but do not assume that faithfulness alone will result in professional success.

Strive to become a true expert in some phase of your area of interest. Superb disciplinary competence is always marketable and can carry you through many difficult times.

“To thine own self be true... thou canst not then be false to any man.” Thus learn to be independent and be able to disagree without being disagreeable.

Give marriage and family high priority; and if you remain single through no fault of your own, find ways to nurture kindred and others. Remember that some things are eternal and celestial in nature. These matter the most even though we must live in and cope with a material world.

Carefully and prayerfully plan your days, months, and years ahead; but do not be surprised nor disturbed if only half or less of your plans come to the expected fruition. In addition to limitations of mortal forecasting and implementing, God and your hidden destiny may intervene in unexpected ways.

Remember that to be learned is good if you follow the counsels of God and his prophets.

Find ways to respectfully and humbly apply your expertise to needs and problems in the Church and other similarly altruistic organizations designed for human betterment.
PUBLISHING IN PSYCHOLOGY: AN OVERVIEW FOR UNDERGRADUATE STUDENTS

This article relies on the expertise of current and former journal editors in order to introduce undergraduate students to important issues relating to publishing in academic psychology journals. These experts have stressed the need for psychology researchers to submit manuscripts that meaningfully contribute to the discipline and are situated within the context of previous research. The importance for undergraduate students of seeking coauthorships with faculty and becoming familiar with appropriate writing style and submission guidelines of the different psychology journals is also stressed. An appendix is included with a list of publication helps for undergraduate psychology students.

Publishing in academic journals is central to advancement in psychology, whether the advancement of knowledge or the advancement of a researcher's prestige. Indeed, those who warn, "Publish or perish!" often prophesy correctly, as many psychology researchers "must publish to obtain a desired position, must publish to retain their position, and often have to publish to advance in the position" (Kupfersmid & Wonderly, 1994, p. 8). Due to the premium placed on publishing, a researcher's ability to frequently publish in academic journals is often considered his or her most valuable asset.

Likewise, publication for an undergraduate student is quite impressive, as well as an invaluable learning experience. In particular, students interested in graduate school and academic careers in psychology would be wise to learn about publishing in academic journals as soon as possible. The purpose of this article is to introduce a few basics of publishing in psychology. I will first demonstrate that researchers contribute to the discipline only through publishing meaningful articles. I will then discuss two ways that undergraduate students can begin working toward publication: collaborating with faculty members and becoming familiar with the content and policies of academic journals. In addition, I have provided an appendix that contains a list of publication helps for undergraduates who are interested in publishing in psychology.

Having Something Meaningful to Say

According to several editors of psychology journals, the primary reason a manuscript is rejected is its failure to meaningfully contribute to the discipline. According to Allan Wagner, former editor of the Journal of Experimental Psychology: Animal Behavior Processes, "By far the most common reason for rejection of papers is lack of substance. . . . If [a researcher's] work represents a genuine contribution, then [editors] will often bend over backward to help the author make the paper acceptable for publication" (as qtd. in Sternberg, 1988, pp. 186–187). Scott Lilienfeld, founding editor of Scientific Review of Mental Health Practice, . . .
agrees, stating that one of the biggest mistakes researchers make is failing to make it clear why their studies are important (Tamashiro, 2003).

One way researchers do not clearly convey the importance of their research is by failing to situate it within the context of previous research. According to BYU’s Brent Slife (personal communication, September 30, 2003), this failure was one of the most common mistakes he saw as a journal editor. When researchers do not place their findings or theories within a proper contextual framework, they fail to advance knowledge in the discipline. The reason for this is logical: if an article does not relate to other research, then it does not relate to other researchers, and consequently, nobody cares. This lack of contextual support is often the result of the failure to understand how, or to what extent, one’s research relates to the previous or potential work of others. As a result, authors often do not realize the redundancy of their studies or are unable to make a strong case for their relevance.

Because such a premium is placed on publication, however, many articles published in psychology journals do not meaningfully contribute to the discipline. Just because a researcher is able to publish an article does not mean he or she has something meaningful to say. In The Psychologist’s Companion: A Guide to Scientific Writing for Students and Researchers, Yale University’s Robert Sternberg (1988)—a prominent expert on writing in psychology—reviewed Tulving and Madigan’s 1970 study, in which less than 10% of articles from a sample of 540 publications were classified as “worthwhile,” according to their “contribution to knowledge” (pp. 166–167). This small percentage, according to Tulving and Madigan, “[carried] the burden of continuous progress in [the] field, by clarifying existing problems, opening up new areas of investigation, and providing titillating glimpses into the unknown” (p. 167). The remaining 90%, however, were classified as “run of the mill” or “utterly inconsequential” (p. 166). These articles either had no bearing on future research or were largely redundant. All of them, Tulving and Madigan predicted, would “fall into oblivion” within one year (pp. 166–167).

Slife agrees that many psychologists publish articles that are not particularly meaningful. He said that many researchers spend their entire careers constructing “pot boilers,” focusing on what is publishable (where the “pot” is hot) rather than on what they can meaningfully contribute. Slife, who serves on the editorial boards of four major psychology journals, said he believes that psychologists will better advance the discipline by being directed by “a strong sense of mission” than a mere desire to add tally marks to a curriculum vita.

Although it may be some time before undergraduate students become immersed in publishing, it is important for them to understand that many published journal articles have little or no impact on others in the discipline. Understanding this, students would be wise to develop a sense of what constitutes a meaningful contribution to the discipline.

Working with Faculty

To an undergraduate student, publishing something relevant in a highbrow academic journal must seem a daunting task. According to Slife, the key for undergraduates is to coauthor a paper with a faculty mentor, as undergraduate students almost never publish papers in academic journals on their own. Slife also recommends that undergraduates “tap into research programs already available,” rather than pursuing their own ideas. Professors are unlikely to sponsor an undergraduate’s ideas because it takes too much time for them to “get up to speed on the literature” for the particular area of the student’s interest. It would also require the student to conduct exhaustive research in order to produce an effective proposal. It is far more feasible for a professor to work with undergraduates on research projects in which the professor is already involved or interested. Slife,
who has published over 120 journal articles and books (as of 2003), said that he coauthored about one-fourth to one-third of his published articles with students and that he is currently working with several students on different projects for which they will receive authorship credit with him (personal communication, September 30, 2003).

In addition, working with faculty members can help undergraduate students frame their research within the proper context. A faculty member's expertise and experience can help students learn ways this is accomplished. (Refer to Appendix for more information about working with faculty at BYU.)

**Becoming Familiar with Academic Journals**

In addition to working with faculty, undergraduate students should become familiar with academic journals. Reading articles in academic journals will help students become familiar with a given journal's prestige, content, and publication guidelines. These aspects are each important considerations in selecting which journal to submit manuscripts to.

In selecting an appropriate journal, researchers often have to weigh a journal's prestige with its likelihood of publishing their articles. Proper selection is important because, according to American Psychological Association (APA) guidelines, a manuscript may be sent to only one journal at a time (Sternberg, 1988). According to Sternberg (1988), "journals vary widely in quality. Some journals publish papers that do little more than fill up journal space; other journals publish only outstanding contributions to the literature" (p. 184). A study's prestige often correlates with the prestige of the journal in which it is published; therefore, it is important for researchers to match their manuscript with a journal of comparable quality. Because rejection rates are much higher in top journals, researchers will be more likely to publish articles in lower quality journals, but their articles will tend to be considered as less prestigious and meaningful (Tamashiro, 2003). Sternberg (1988) recommends sending a manuscript to one's first-choice journal while having alternative choices in mind in case of rejection. As undergraduate students gain familiarity with journals, they can begin to estimate the quality of manuscript that is required for a particular journal's consideration.

In addition, as students read academic journals, they will inevitably become familiar with their content and editorial guidelines. This is important because every editorial board limits the types of manuscripts they accept according to a particular focus regarding appropriate topics and methodologies (Sternberg, 1988). Many articles are rejected simply because they are not appropriate for a journal's audience. Some journals (e.g., *American Psychologist* and *Psychological Science*), for example, have general readerships and therefore contain articles that are of interest to psychologists in general. Consequently, the editorial board of *Psychological Science* gives preference to "articles that are deemed to be of general theoretical significance or of broad interest across specialties of psychology and related fields, and that are written to be intelligible to a wide range of readers" (Association for Psychological Science, 2006, inside back cover, emphases added).

According to *Psychological Science* 's editor, James Cutting, manuscripts are often rejected merely because they are too specialized (Tamashiro, 2003). An example of a more specialized journal is the *Journal of Family Psychology*, an APA journal "devoted to the study of the family system from multiple perspectives and to the application of psychological methods to advance knowledge related to family research, intervention, and policy" (American Psychological Association, 2006b). Its audience is largely composed of professionals, particularly APA members who specialize in family research and therapy.

In addition to considering the scope of a journal's audience, researchers should consider
length restrictions, publication lag, and authorship restrictions (Sternberg, 1988). Most journals will publish submission guidelines regarding these elements in each issue, as well as on their Web sites. For example, one learns from reading the submission guidelines of American Psychologist that manuscripts “may not exceed 35 double-spaced pages in length, including the cover page, abstract, references, tables, and figures” (American Psychological Association, 2006a). Publication lag refers to the average length of time a publisher waits to publish an accepted manuscript; this length of time might make a difference in selecting an appropriate journal. It is also important for researchers to be aware of authorship restrictions in certain journals; for example, some journals publish only articles written by members of a certain organization (Sternberg, 1988). (Refer to Appendix for more information on becoming familiar with academic journals.)

In conclusion, it is important for undergraduate students to understand that the discipline of psychology is largely centered on publishing in academic journals. Researchers best contribute to the discipline when they submit findings or theories that are meaningful to other researchers. Undergraduates are much more likely to publish by working with faculty members on previously established research projects. They will also improve their knowledge of submission guidelines and appropriate writing styles as they read and become familiar with the academic journals.

References


Appendix

Publication Helps for BYU Undergraduate Psychology Students

The following list includes helps for undergraduate students interested in publishing in psychology. Some helps are unique to BYU students; others are for students or writers in general.

• As discussed above, the key to publishing as an undergraduate is to seek coauthorships with a professor in an area of his or her expertise and interest. To review faculty interests for BYU’s psychology department, consult the department’s Web page (http://psychology.byu.edu/people/faculty.html) or visit the undergraduate coordinator at Psych Central (1150V SWKT).

• It will be difficult to publish without completing the core skill courses in psychology (Psychology 301, 302, and 304). To maximize your chances of publishing while an undergraduate, be sure to complete these courses early.

• Both the American Psychological Association (APA) and the Association for Psychological Science (APS) inform student affiliates of publication opportunities, some of which
are exclusive to undergraduates. Annual membership for undergraduates costs $27 for APA and $35 for APS. For more information, consult www.apa.org (APA) and www.psychologicalscience.org (APS).

- Learn about publication standards and the process of submitting to psychology journals by visiting their Web sites. Most mainstream psychology journals are affiliated with APA; information for each can be accessed at APAs home page (http://www.apa.org). Likewise, the APS home page provides information concerning its journals (http://www.psychologicalscience.org).

- Psi Chi, the official student honor society in psychology, provides several opportunities for undergraduate publication. BYU’s chapter generally has membership drives at the beginning of each semester. Cost for lifetime membership is $45. Membership is limited to those who meet certain academic requirements. For more information, visit the Web site of the BYU chapter (http://clubs.byu.edu/psichi/index.html) or the national organization (http://www.psichi.org) or keep an eye on the Psi Chi bulletin board (main level, SWKT).

- For students serious about publishing, understanding APA writing style (the dominant style used in psychology journals) is especially helpful. Regularly refer to, or perhaps purchase, a copy of the Publication Manual, 5th edition. For a quick overview, a handout concerning APA documentation is available at the BYU Writing Center (B106 JFSB). The handout is available online: http://english.byu.edu/writingcenter/popups/apa.pdf. At the center, copies of the Publication Manual are also available for perusal.

- A great place to begin publishing is intuition: BYU Undergraduate Journal of Psychology. In addition, joining the Intuition staff can be a great way to become familiar with publishing standards and skills in psychology. For more information, visit the Intuition Web site (http://intuition.byu.edu/index.html) or send an e-mail to the editors (byupsychjournal@gmail.com). Credit is available for staff members.
Tokenism in Non-Native English Speakers: Does It Affect Their English Testing Abilities?

This study was conducted with 16 international non-native English speaking undergraduate students to examine the impact of tokenism on such students' test-taking performance. The students were divided into two testing conditions. The experimental group was given a modified GRE verbal ability test in a room with three native English speaking confederates. The control group was given the same test in the presence of fellow international students. The purpose was to compare test results of the two groups, taking into account participants' TOEFL scores, age, and length of stay in the United States. A significant difference was found between the experimental and control groups when accounting for length of stay in the United States.

Over the past decade the number of international students enrolled in colleges and universities in the United States has increased dramatically (Bollag, 2004). Currently, over 572,509 international students are attending American colleges and universities (Bollag, 2004), with approximately 2,000 at Brigham Young University (Light, 2004). Researchers have acknowledged the important relationship between English proficiency and academic success for international students (Prieto, 1995) but have failed to recognize other possible obstacles to the academic success of international students. The purpose of this study was to identify and gain a better understanding of one possible obstacle, the negative impact of token status on international students' test-taking performance.

Tokenism, according to Inzlicht and Ben-Zeev (2000), is a situational phenomenon in which an individual representing a minority group experiences cognitive deficits in performance, even when treated the same as majority group members. Recent research regarding tokenism suggests that when token minority students are placed in a group in which their token status is salient, the students' achievement levels suffer (Steele & Aronson, 1995). This occurred even when the students were high achievers and did not believe in a minority stereotype (Steele & Aronson, 1995). Furthermore, even when students were stereotyped with high achievement in a given domain, such as high performance in mathematics among Asian students, token minority status still interfered with intellectual performance (Spencer, Steele, & Quinn, 1999). According to Steele (1997), the cause of the underperformance may have been a fear of confirming the stereotype and being judged as a member of a stereotyped minority group.

This study is based on a recent study of tokenism conducted by Inzlicht and Ben-Zeev (2000), regarding women and their math test performances in predominantly male settings. They predicted that because females were outnumbered by males in the test-taking situation, they would score significantly lower in mathematics performance. In fact, females in the co-ed test setting did score significantly lower...
on the math portion compared with females in the single-sex (all female) test setting. This result was found even without reminding participants of the stereotype that men typically perform better than females in mathematics.

Furthermore, token stereotypes affecting academic achievement are also found among international students in American universities. Recent studies of stereotyping found that in colleges and universities, token symbolism has been identified with international students in the United States—particularly concerning language proficiency (Spencer-Rogers, 2001). Some of the stereotypical views for token international students are that token minority students “do not speak English well” or are “maladjusted” (Spencer-Rogers, 2001). Leki (2002) reported that international students often feel negative stereotyping in scholastic group settings with native English speaking students.

Thus, international students being tested amid native English speakers in classroom settings may score significantly lower than their native English speaking peers, due to awareness of their token minority status. While other studies investigating tokenism have predominantly assessed women and math performance (Inzlicht & Ben-Zeev, 2000; Lord & Saenz, 1985; Saenz, 1994; Saenz & Lord, 1989), this study investigated the effects of token minority status among international students assessed on verbal performance.

The objective of this study was to examine whether international students’ test performance suffers significantly due to students’ token minority status. This objective was investigated by placing international students in a test-taking environment in which they were outnumbered (4 to 1) by native English speaking students. We hypothesize that international students who are tested among native English speaking confederates will score significantly lower on a verbal exam when compared with international students taking the same test among other international students.

Method

Participants
Participants were 16 undergraduate, non-native English speaking international students, 5 male and 11 female, attending Brigham Young University. They were recruited from undergraduate psychology courses and received extra credit for their participation. Native tongues varied considerably (4 Spanish, 1 Romanian, 2 Portuguese, 2 Tai, 2 German, 1 Fijian, 1 Romanian and Russian, 1 Mongolian, 1 Mandarin, and 1 Korean).

Materials
An anonymous consent form was used (i.e., the subjects did not sign it; they showed their consent by handing in the questionnaires).

The verbal test consisted of 23 multiple-choice questions similar to those on the Graduate Record Exam (GRE)—14 vocabulary definition questions, 3 reading comprehension questions, and 6 word analogy questions. For each question, participants were asked to circle the letter representing the “best answer.” Pencils were provided.

Two questionnaires were also administered (see Appendices 1 and 2). In the first questionnaire (Appendix 1), participants ranked themselves on a five-point Likert-type scale (ranging from strongly agree to strongly disagree) regarding their level of comfort during the test (e.g., “I feel that I could have scored better had I taken this test alone”). The other questionnaire (Appendix 2) surveyed demographic information such as age, Test of English as a Foreign Language (TOEFL) scores, and the length of time they have lived in the United States.

Procedure
Based upon their availability to take the test on a particular day and at a specific time, half of the participants were assigned to the experimental condition and the other half were assigned to the control condition. All participants sat at tables.
and took the test in either a conference room or a classroom. Most participants were uninterrupted during the test-taking time, though this was not completely controlled for (several students entered the rooms while the test was being administered, despite a sign asking them not to). There was ambient lighting in each of the test rooms.

For the experimental condition, each of the eight subjects was tested separately. When they entered the testing room, there were three Caucasian confederates waiting and talking together. The participants were asked to take a seat in the fourth chair and were instructed to read the consent form. Instructions were then given regarding testing procedures. Participants were given 25 minutes to complete the test, with a five-minute warning. After testing, participants filled out both of the aforementioned questionnaires. Following, the researcher read a prepared statement to debrief participants.

For the control condition, eight participants were divided into two groups of four and both groups took the test without the presence of confederates. All other testing procedures, however, were similar to those of the experimental group.

**Statistical Procedures**

Initially, mean test scores for the experimental and control group (\( N = 16; n_1 = 8, n_2 = 8 \)) were compared using independent t-tests. These two groups were each divided into subgroups (\( n_{1a} = 4, n_{1b} = 4, n_{2a} = 4, n_{2b} = 4 \)) three separate times, based upon individual TOEFL scores, age, and length of stay in the United States (see Figure 1). Separate ANOVAs were performed for each cluster of subgroups. Statistical powers for this study were low because of the small sample size, 16 participants (see Results for exact \( \eta^2 \) values).

**Results**

Initial t-tests comparing the difference between test scores of the control and experimental groups (\( n_1 = 8, n_2 = 8 \)) yielded no significant results. The two ANOVAs on subgroups based upon participants’ age and TOEFL scores also yielded no significant results. However, the ANOVA on subgroups based on subjects’ length of stay in the United States showed significant interaction between length of stay in the United States and test scores (\( p = 0.02, df = 6, \eta^2 = 0.11 \)) (see Figure 2). The same ANOVA also showed a slightly less than significant interaction (\( p = 0.07, df = 6, \eta^2 = 0.07 \)) between test scores of the subgroups that were tested without confederates present (see Figure 3).

**Discussion**

Based on tokenism theory, which states that token minority status results in cognitive deficits (Lord & Saenz, 1985; Saenz, 1994; Saenz & Lord, 1989), it was predicted that the test scores for international students tested with other international students would be significantly higher than those of international students tested in the presence of native English speaking confederates. T-tests revealed no significant differences. The results do indicate, however, that participants who had lived in the United States for a shorter period of time scored significantly higher on the verbal exam, when tested with other international students, compared with their counterparts who were tested in the presence of native English speaking confederates. However, no significant differences were found between the scores of international students who had lived in the United States for a longer period of time.

Based on our review of literature, this is the only study to apply tokenism theory to international students who speak English as a second language. Other studies investigating tokenism predominantly draw upon women and native English speaking minorities as participants (Inzlicht & Ben-Zeev, 2000; Lord & Saenz, 1985; Saenz, 1994; Saenz & Lord, 1989). Furthermore, the interaction regarding international students’ length of time.
Figure 1: Test Organization Diagram showing the formation of groups used in the statistical analysis of the data.
Figure 2: Graph showing the mean number of correct answers obtained on a verbal test by subjects who had a short length of stay in the United States, divided by those assigned to test with confederates present (the experimental group) and those assigned to test without confederates (the control group). Error bars represent ± one standard deviation. These results were significant at the \( p = 0.02 \) level with \( df = 6 \).

Figure 3: Graph showing the mean number of correct answers obtained on a verbal test by subjects who tested without the presence of confederates, divided by long or short length of stay in the United States. Error bars represent ± one standard deviation. These results were not significant.
lived in the United States and their test scores support findings of other researchers of tokenism—suggesting that the token status of an individual does facilitate cognitive deficits in performance.

One confound of the study was the frequent interruption by students entering the conference room or classroom while the test was being administered. These interruptions may have affected test scores—multiple subjects complaining that the interruptions made it more difficult for them to concentrate. Another problem was that a few of the subjects saw through the deceptive claim that they were being tested on verbal competency. After the examination, these subjects, who were tested in the presence of confederates, reported that they suspected that the true purpose was to see how they would perform when tested with native English speakers. The effects of this awareness varied across participants’ questionnaire responses, with participants reporting both increased and decreased anxiety; this unforeseen confound could have possibly affected the participants’ test scores. In future research, better controls should be used in order to conceal the deceptive nature of the study.

This study had several other limitations. One such limitation was that the conditions were not randomly assigned. International students were recruited and assigned to a testing group based on their availability to take the test on a particular day and at a specific time. Furthermore, the test administration procedure was not double-blind. The test administrator was aware of international students' token status, which may have caused administrative bias (e.g., greater length of eye contact with the participant, compared with confederates, during the instruction period of test administration).

Another major limitation was the small sample size. Only 16 students participated in the study—one international student in the non-confederate testing situation and 8 in the confederate testing situation. For future research, a larger sample size will be needed to increase statistical power and therefore accurately determine if token status facilitates cognitive deficits specific to verbal performance.

Finally, subjects tested in the presence of confederates may not have been aware that they were in the presence of native English speakers and may thus have been ignorant of their token status. It is possible that the confederates were mistaken for Caucasian international students from non-English speaking countries. More salient “getting-to-know-you” conversation amongst confederates, when the participant first entered the room, may have convinced participants that the other test takers were native English speakers.

It is also possible that other theoretical viewpoints may be useful in explaining the results of this study. Stereotype threat and distinctiveness theory may be especially useful. Stereotype threat is a phenomenon that occurs when stereotyped individuals are reminded of their negative stereotype and the possibility of the stereotype’s validity (Aronson et al., 1999; Aronson, Quinn, & Spencer, 1998; Spencer, Steele, & Quinn, 1999; Steele, 1997; Steele & Aronson, 1995). It is possible that international students experienced stereotype threat during test administration because of common stereotypes suggesting that immigrants and foreigners do not perform as well as natives in various situations (Spencer-Rogers, 2001).

Distinctiveness theory states that being a member of a minority group can lead to a sense of group identity which then becomes part of the working self-concept (Abrams, Thomas, & Hogg, 1990; McGuire, McGuire, Child, & Fujioka, 1978; McGuire, McGuire, & Winton, 1979; McGuire & Padawer-Singer, 1976). Thus, in this study, participants in the experimental condition may have identified themselves as being distinct from the confederates, incorporating that distinction into their self-concept while taking the test, ultimately affecting their performance. Therefore, though this study’s results support token theory, we suggest that future research investigate other theoretical explanations for similar studies.
References


Appendix 1: Self-Report Ratings

Please read each statement and decide how much you agree or disagree. Please circle only one of the five choices for each statement.

SA = Strongly Agree; A = Agree; N = Neutral; D = Disagree; SD = Strongly Disagree

1. This examination made me anxious.
SA A N D SD

2. I think I could do much better on this test if I could take it alone.
SA A N D SD

3. During the test, people around me made me nervous.
SA A N D SD

4. Thoughts of doing poorly interfere with my performance on the test.
SA A N D SD

5. Even I have been taking college courses in English, I feel very anxious about taking this test.
SA A N D SD

6. While I was taking this test, I found my hands or arms trembling.
SA A N D SD

7. I was relaxed while I was taking this test.
SA A N D SD

8. While I am taking this exam, I found myself thinking of how much brighter the other students are than I am.
SA A N D SD

9. Other people in the room seem much smarter than me.
SA A N D SD

10. I got to feeling panicky when I was taking this exam.
SA A N D SD

11. During this test, I found myself thinking of the consequences of performing poorly.
SA A N D SD

12. I felt so tense and my stomach got upset.
SA A N D SD

13. When taking this test, my emotional feelings interfere with my performance.
SA A N D SD

14. During this test, I got so nervous that I could not perform as much as I wanted.
SA A N D SD

15. While I was taking this test, I thought about how awful my verbal ability is.
SA A N D SD

16. I felt unintelligent while I was taking this test.
SA A N D SD
Appendix 2: Demographic Questionnaire

Age (please fill in the blank below)

Sex (please circle one)
- Male
- Female

TOEFL Score (please fill in the blank below: estimate if necessary)

SAT or ACT Verbal Score (please fill in the blank below: estimate if necessary)

Native Language (please fill in the blank below)

Length of Stay in the United States (to date)
- Year(s)
- Month(s)

Authors’ Note

The authors wish to acknowledge Dr. Niwako Yamawaki for her indispensable assistance in the formulation and execution of this research endeavor.
THE EFFECTS OF SEMANTIC RELATIONSHIPS ON
THE IRRELEVANT SOUND EFFECT

This study investigated whether the degree of semantic relationship between audible distracter words and visually displayed to-be-remembered words impact short-term memory recall. Semantic relationship was defined as the level of synonymity between the two categories of words. Participants were divided into a control group receiving no distracters, a group in which distracter words and to-be-remembered words were closely synonymous, and a group in which distracter words and to-be-remembered words were loosely synonymous. The results indicate that semantic relationships have no impact on the number of words correctly recalled. However, closely synonymous relationships caused more replacement errors than silence or loosely synonymous relationships. One possible explanation is that auditory stimuli are more readily perceived than visual stimuli when synonymity is high.

Several types of distracters can inhibit the formation or recall of memories. While there is more time for interference to occur with long-term memories than short-term memories, short-term memory (STM) can still be disrupted. One type of distracter is external auditory stimuli. Sounds not associated with a task involving short-term memory negatively affect recall. This effect is termed the irrelevant sound effect (ISE; Colle & Welsh, 1976).

It is uncertain which features of auditory stimuli cause the greatest impact on the level of disruption in the ISE. Some studies indicate that acoustic properties are the key features of the ISE (Salamé & Baddeley, 1982; Tremblay, Macken, & Jones, 2000). Acoustic properties, such as the complexity and non-continuity of sounds, cause significant disturbance, with non-continuous and more complex sounds being more disturbing (Jones, Macken, & Murray, 1993). For instance, a continuous tone (such as a siren) would be less disturbing than a pulsating alarm. Also, a complex multi-toned siren would be more distracting than a siren with only a single tone. However, auditory stimuli experienced during memory recall include more than tones, music, and noise. More specific than the acoustic properties of general sounds are the acoustic properties of human speech. The acoustic properties within human speech are referred to as phonological properties. LeComte and Shaibe (1997) reported that there was no phonological effect, meaning that acoustic properties within irrelevant speech do not affect recall. These studies suggest that human speech is not a significant cause of disruption (Buchner, Irmin, & Erdfelder, 1996).

While the previously cited studies indicate that the acoustic properties of speech do not impact STM recall, some research posits that semantic properties of the irrelevant sound are significant in the ISE (Neely & LeComte, 1999). Semantic properties are those properties associated with a word's meaning. Tests originally claiming that semantics had no effect lacked statistical power because of low subject counts (Buchner, Irmin, & Erdfelder, 1996). Buchner, Irmin, and Erdfelder (1996) conceptually replicated Salamé and Baddeley's (1982) study while including more
subjects. They also concluded that semantics had no effect. However, these tests were designed to measure the semantic effects of digits, not words. In experiments testing the effects of words with more semantic meaning than digits, results indicated that the semantic relatedness of distracter words to the to-be-remembered (TBR) words was a significant factor in disruption of STM (Neely & LeCompte, 1999; Oswald, Triebay, & Jones, 2000). There is a difference between how the sounds involved within human speech distract and how the meanings involved with those sounds distract.

Only a few studies have investigated in any depth the significance of semantics involved with the ISE. Neely and LeCompte (1999) used distracter words that had related meaning to the TBR words. During trials where the distracter words were related to the TBR words, both sets of words came from the same general category of word types (e.g., all words were types of fruits; Neely & LeCompte, 1999). Other research reports that the valency of distracter words has a negative effect on recall, particularly negatively valenced words (Buchner, 2005). It has also been found that the frequency of repeated words within a sentence proportionally increases the effect of the ISE (Buchner, 2005). These findings provide insight into how the semantic aspects of words actually impact recall. However, much is still to be learned about how related words impact recall.

Studies based on Buchner’s (2005) work involving valence and word frequency and modeled after Neely and LeCompte’s work (1999) might further the understanding of semantic elements in the ISE. Neely and LeCompte’s results indicate that words within the same category negatively affect recall more than distracter words from different categories. There is much more to semantics than nouns within the same category. Elements of speech that are more abstract, such as the synonymous conceptual meanings of words, may also negatively affect recall in a similar manner to Neely and LeCompte’s word groupings.

Much research has been done to establish whether or not semantics affect recall ability. What remains is the need to understand how and to what degree the effect occurs. Many aspects of semantics may impact recall. Our study examined the impact of words with similar meanings. Previous research identified a difference in recall when distracter words were in the same word grouping as the TBR words (Neely & LeCompte, 1999), but all of the words were distinct in meaning. Our study extends Neely’s work to examine differences in meaning within word groups. Specifically, we examined the effect of distracter words that were synonyms of the TBR words.

It was hypothesized that irrelevant sounds will generally inhibit recall of visually presented TBR words. Furthermore, auditory distracter words that are close synonyms of visually displayed words will produce fewer correctly recalled words than auditory distracter words that are loose synonyms of visually presented words.

Method

Participants

Sixty-seven college students participated in the study. These students were recruited from psychology classes at Brigham Young University and were all fluent in English. In most situations the students received extra credit for their participation. Researchers went to introductory psychology classrooms and asked for volunteers to participate in a memory test scheduled for a classroom location on campus at various times.

Participants were randomly assigned without replacement to three groups by asking them to choose a coin out of a sack. The year on each coin determined the groups: 2001 for group 1, 2002 for group 2, or 2003 for group 3.

Materials

The experiment used an Apple PowerBook G4 (model 6, 4) laptop computer, Sony MDR-V600
headphones, a pen, and 3" x 5" cards. Microsoft PowerPoint 2004 for Macintosh was used to flash words on the computer screen, and Garage Band v 1.1.0126 was used to paste in the auditory background words. A one-second time delay was used in PowerPoint to flash the words on the screen one at a time (see Table 1).

Design and Procedure

Participants were asked to sit down at the computer, which was placed on a desk at the front of a classroom. They were given a 3" x 5" card and a pen and were told to put the headphones on. They were instructed by a researcher to push the space bar when they were ready to start and that the computer would instruct them on the rest of the procedure. The researchers made sure not to stare or look at participants while they were taking the test. After the participants completed their tests, one of the researchers asked them if they had any questions. If they did, their questions were answered as their card was collected. The participants were debriefed, receiving a description of the nature and purpose of the experiment. Participants were told that their performance appeared to be normal and that they had adequately performed the task. They were thanked for their participation and excused.

Recall Test

The timing, words used, instructions, order of the slides, and the presence or absence of sounds are summarized below.

Slide #1: A blank screen displays until participant presses the space bar.

Slide #2: Slide displays: “Listen to the instructions. Press space bar when you’re ready.”

Participant hears: “Thank you for your participation in this study. This is not a test of skill or intelligence, so please make yourselves comfortable and relax.

“A list of 10 words will be presented to you. Each word will be displayed for approximately one second. Please memorize each word as it is displayed, ignore any word or sounds you hear, and only pay attention to the words on the screen. Once all the words have been displayed, 10 seconds will elapse before you are prompted to record what words you can remember in any order on the note card provided. You will be given 30 seconds to record. We will begin now.”

(Participant presses the space bar to start the word list.)

Slides #3–#12: (Each visual word is displayed for 1 second for all groups. Auditory words are heard simultaneously with the visual words displayed except for in group 1, which is the control group. Close synonyms are heard in group 2 and loose synonyms in group 3. See Table 1 for the word lists.)

Table 1

Visual words
(same for all groups):

<table>
<thead>
<tr>
<th>house</th>
<th>engine</th>
<th>book</th>
<th>human</th>
<th>student</th>
</tr>
</thead>
<tbody>
<tr>
<td>country</td>
<td>wish</td>
<td>parent</td>
<td>field</td>
<td>school</td>
</tr>
</tbody>
</table>

Auditory distracter words in group 2
(closely related):

<table>
<thead>
<tr>
<th>home</th>
<th>motor</th>
<th>text</th>
<th>person</th>
<th>learner</th>
</tr>
</thead>
<tbody>
<tr>
<td>nation</td>
<td>hope</td>
<td>father</td>
<td>pasture</td>
<td>college</td>
</tr>
</tbody>
</table>

Auditory distracter words in group 3
(loosely related):

<table>
<thead>
<tr>
<th>abode</th>
<th>machine</th>
<th>tome</th>
<th>creature</th>
<th>novice</th>
</tr>
</thead>
<tbody>
<tr>
<td>terrain</td>
<td>fancy</td>
<td>author</td>
<td>patch</td>
<td>system</td>
</tr>
</tbody>
</table>

Slide #13: A blank slide appears for 10 seconds before participants are prompted to record words.
Slide #14: Slide displays: “Record answers now.” Slide changes after 30 seconds. (Participant records answers with the pen and note card provided.)

Slide #15: Slide displays: “Stop recording now.” Participant hears: “Stop recording now.”

**Data Analysis**

The number of words correctly recalled and written down by subjects in each of the three groups was recorded. A one-way ANOVA was run on these results. In examining the 3” x 5” cards we discovered that many participants in group 2 (close synonyms) and only one in group 3 (loose synonyms) recorded distracter words. The distracter words were also analyzed. A one-way ANOVA in a secondary analysis compared the number of distracter words that replaced TBR words.

**Results**

The number of correctly recalled words from the three groups was recorded, as was the number of distracter words written down in each of the three groups. A one-way ANOVA was used to determine whether the differences were statistically significant. SPSS was used for all the statistical analyses.

**Analysis of Correctly Recalled and Recorded Words**

When correctly recalled words (see Table 2) were compared across the three groups, it was found that the differences were statistically significant \( F(2, 64) = 8.95 \) \( (p < .0001) \). A Tukey post-hoc test revealed that the mean differences were significant \( (p < .05) \) between groups 1 and 2 and between groups 1 and 3, but not groups 2 and 3 \( (M = 6.33, M = 4.8, \) and \( M = 5.35 \) for groups 1, 2, and 3 respectively).

**Analysis of Replacement of TBR Words with Auditory Distracter Words during Recall**

Group 1 replaced none of the TBR words with the auditory distracter words during recall, while...
group 2 had a mean of 0.36 words per participant and group 3 had a mean of 0.05 words per participant (see Table 3). An F-test showed that there were statistically significant differences ($p < .05$) between groups 2 and 3 in the number of words replaced (see Table 3). Tukey post-hoc tests revealed mean differences between group 1 and 3, and 2 and 3. This indicates that the synonyms had a small but significant effect on the replacement of TBR words with distracter words.

**Discussion**

Our first hypothesis was that any auditory distracter words would be more disruptive than silence. Our second hypothesis was that distracter words that are more synonymous with TBR words would be more disruptive of recall than words that are less synonymous. That is, fewer words would be recalled correctly when the semantic relationship between words was closer. The finding that groups 2 and 3 recalled fewer words correctly than group 1 did support our first hypothesis. Using only the number of correctly remembered words, the difference between the two groups who heard different distracter words was not significant. Thus, the results do not support the second hypothesis.

In a secondary analysis, we found that the differences between replacement error rates between groups were significant. A replacement error was defined as the recall of a distracter word in place of the recall of a visually presented word. Following a one-way ANOVA, a Tukey post-hoc test indicated that the differences between groups 2 and 1 and between groups 2 and 3 were significant, with group 2 having the most replacement errors. The difference between group 3 with loosely synonymous distracter words and the control group 1 with no irrelevant speech distracters was not significant.

These findings from the secondary analysis indicate that, when replacements of words occur, closely synonymous relationships impact replacement errors more heavily. Loosely synonymous relationships have no more bearing on replacement errors than silence does. The results of the secondary analysis extend beyond our original second hypothesis. Specifically, the second hypothesis was that fewer words would be recalled when there was a closer relationship between words. However, in a more general form, the second hypothesis was that closer synonymous relationships between distracter and TBR words would be more disruptive of memory recall.

The ISE not only is impacted by semantics but is impacted in a subtler manner than previously indicated. Taken together, the results indicate that varying the levels of semantic relationship between visually presented words and auditorily presented words has no impact on correct recall of the visually displayed words. However, the recall errors when the visually presented and audibly presented words were more closely related were more frequent than when the relationship was looser. Participants were less likely to distinguish between what was seen and what was heard when the two words were more closely related. This difference supports the general hypothesis put forth by Buchner (2005) and Neely and LeCompte (1999), who posited that semantics impact recall. The present study extended Neely and LeCompte’s (1999) test beyond using only words from the same semantic family during each trial to include words from a variety of families within one trial. Also, the distracter words were related to the visually displayed words in a conceptual manner rather than having direct item-family relationships as in Neely and LeCompte’s (1999) test, where words were related in that they were, for example, all fruits. Both the increased complexity of the word list and the looser word associations between visually and auditorily presented words aided in uncovering these subtle influences in distraction that had not been previously expected.

No studies within the literature have addressed the issue of replacement by the distracter word in the ISE. This may be due to the testing design used thus far in studies attempting to identify the
role of semantics in the ISE. Buchner (1996, 2005) used serial recall in his tests. Neely and LeCompte (1999) also used serial recall but were not studying degrees of relatedness. Serial recall tests may not lend themselves to the detection of replacement words. However, free recall tests allow this. The current study showed that closely synonymous relationships between distracter words and TBR words increased the likelihood of the distracter words to be recalled. Free recall tests may make it easier for such replacements to be measured, but whether they will help to parse the differential effects of different levels of semantic relationships on word replacement remains to be determined. To further this study, the reliability of free recall versus serial recall versus multiple-choice tests in detecting these differences should be better established.

One explanation for the replacement in recall is that two semantically related words may tap into a schema in which both words are present. Whichever word is more salient or more highly prioritized within the individual's schema will be recalled. Each schema and prioritization order would differ subjectively. This also assumes an equal-priority balance between modes of perception. That is to say, this explanation assumes that neither the auditory nor visual pathways should have greater access to the schema than the other. Attention to stimuli from both auditory and visual sources being equal, they should receive equal priority in accessing the schematic information. Thus, whichever word of the two presented to the subject is more highly prioritized within the schema at any point in time will be recalled more readily, regardless of which channel the recalled word was presented in.

This putative explanation is unlikely for at least two reasons. As the participants were specifically told to pay no attention to audio stimuli, words stored in memory that matched what had been visually presented should have had greater priority. Also, while personally relevant words may cause attention to be shifted to otherwise irrelevant speech (Buchner, Irmen, & Erdfelder, 1996), the personally neutral words in the present study should not have caused the participants to attend to them when they were instructed to ignore them. These two reasons could be assured by better controlling for the neutrality of the distracter words and determining the extent to which instructions to ignore auditory words were obeyed.

Another possible explanation for the results we have termed the "aural-trust" phenomenon. It is possible that when disparity between related words is low, the brain is more likely to perceive what was heard rather than what was seen. Therefore, when there is a closely synonymous relationship between the words, the auditory presentations are perceived ahead of the visually presented words. Such a mechanism may explain the pattern of replacement that was observed.

In addition, there were four internal limitations within the current study. The first involved the control of semantic influence. In hindsight, more information about the effects of semantic influence on the ISE may have been obtained with a fourth group included in the research design. This fourth group that had received irrelevant speech distracter words with no semantic relationship to the visually presented TBR words would have acted as a secondary control for the effect of semantic relationship on recall; that is, they would have provided a baseline for no semantic relationships.

Another limitation was the word list construction. Synonyms are difficult to rank by order of similarity, especially while simultaneously controlling for word length and phoneme-matching between word pairs. Subjective differences in how synonyms are related to an individual may vary, as already noted in the previous discussion of schemas, causing differences in the words that are close enough in meaning to be replaced during recall. The disparity in word meaning that was a prospective cause of replacement may have varied from person to person. The disparity may also have varied in degree within some word lists, with some pairs in a group being more closely related than other pairs. Perhaps one way to control for a
subjective difference in the perceived synonymity of word pairs would be to ask participants to rate word pairs according to synonymity after the recall task. This rating, if factored into the participant's results, may provide for better control of synonymity differences when examining replacement errors.

The third limitation was that the subjects were not isolated during the tests, allowing for distraction beyond that which was purposely included in the design. While this was controlled for as much as possible, the nature of room scheduling inherent in using the campus made it impossible to isolate each student in order to be undistracted by anything other than distracter words. Participants may have been distracted by background noise other than what the study controlled for. Several participants were observed to look about the room at other participants and researchers during the course of the session. Occasionally external sounds (such as the door being closed loudly or other persons speaking loudly) disrupted an otherwise silent testing environment.

The fourth limitation in the study was that some of the words in the auditory portion of the slideshow were not completely audible due to recording difficulties. Some of the words ended slightly prematurely, causing a minor abnormality in their sound. Three of the words had this defect. This irregular sound of the words may have confounded the semantic effects.

While it is possible that word replacement is the only effect of the ISE connected with semantics, the extent of semantic effects should be more thoroughly probed. Although the present study narrowed the effect of semantics on the ISE more than previous research, it was nevertheless still a broad approach. More detailed analysis of the impact of synonyms must be done by eliminating potential confounds. As our hypothesis was not focused on the replacement of target words by distracter words, research pointed at this issue should employ similar word lists while using free recall tests specifically aimed at discovering how and why words are replaced. This could be done by varying the disparity between synonymous word groups to better describe the continuum of effect. The number of target words used in the test should be examined as well. In this experiment the word list included only 10 words. Most people can remember only seven items, plus or minus two, without any distractions. A longer list may cause the retention of words to drop at different rates depending upon a given distraction. The order in which the words are administered should be analyzed to determine if these two effects override any ISE. Research should also investigate whether there is a threshold for the sound levels at which the ISE occurs.

References


THE INFLUENCE OF PERCEIVED PARENTING STYLES ON THE DEGREE OF ADULT CHILDREN'S ALLOCATION OF PUNISHMENT

Though extensive research has been done investigating parenting styles and the influence that those styles have on children (e.g., Baumrind, 1971; Barry, 1991; Lamborn, Mounts, Steinberg, & Dornbusch, 1991), the influence of parents and their parenting styles on their adult children's assignment of punishment has not been thoroughly examined. To illustrate this relationship, research was conducted measuring the degree and type of punishment that 84 adult participants would allocate for both civic and household offenses. Level of punishment was compared with the perceived parenting styles of the respondents' parents, as measured by Bar's (1991) Parental Authority Questionnaire (PAQ). The adult children of the authoritative fathers issued significantly more severe punishments than their authoritative fathers (M = 4.6 non-authoritative, M = 2.98 non-authoritative).

The influence that various parenting styles have on the behavior and development of their children has been investigated by various researchers (e.g., Grusce & Goodnow, 1994; Lopez, Bonenberger, & Schneider, 2001; Pratt, Arnold, Pratt, & Diesner, 1999). Most current research in parenting practices uses Baumrind's (1971) popular classification system, which categorizes parenting styles according to two dimensions of parental influence: level of expectation (or level of demand upon the child) and level of responsiveness to the child as an individual. Parents who have high expectations and are responsive to their children are classified as authoritative; parents who have high expectations but are not responsive are classified as authoritarian; and parents who have low expectations and are not responsive are classified as permissive. Lastly, parents who are low in both dimensions are considered neglectful or uninvolved (Pratt et al., 1999). This final category is generally considered to be an absence of parenting rather than an implemented "parenting style" and was, therefore, not considered in the present study on parenting styles. Authoritative parenting is typically considered the most effective parenting style for producing healthy, well-adjusted children, whereas authoritarian and permissive parenting have been seen as having the tendency to hinder social and moral development (Lopez et al., 2001; Pratt et al., 1999).

These parenting styles have been linked to the development of children's moral reasoning (Lopez et al., 2001). Moral reasoning is the process of making decisions concerning right and wrong based on social norms and ethical principles. The development of moral reasoning is thought to reach its pinnacle when a person is able to both view a situation from the perspective of another and base decisions on universal principles (Craim, 1985).

The first aspect of moral reasoning, the ability to take the perspective of another person and understand that person's feelings or intentions, is called empathy. In a study examining how different age levels would assign punishment for various actions, Helwig, Zelazo, and Wilson (2001) found that
children who have reached a higher stage of moral development tend to take into account intentions of an individual, rather than just the consequences of the individual's actions. Paying attention to intentions suggests an increased level of empathy or an ability to better identify with others.

Empathy is also closely related to the principle of distributive justice and is influential in its implementation. Distributive justice is the principle through which individuals seek a correlation between rewards and some level of deservingness on the part of a recipient (Hoffman, 2000). Distributive justice generally applies to allocating rewards, such as the distribution of points to individual students in a group who have worked collaboratively on a project. However, the same idea may be used in considering the designation of a specific punishment as a consequence of wrongdoing.

According to Hoffman (2000), empathetic feelings motivate people to treat others more mercifully. Empathy may encourage an individual to adhere less rigidly to a universal code or concept of justice as he or she makes judgments (see Batson, Klein, Highberger, & Shaw, 1995). Thus, an individual who feels empathy for another tends to show preferential treatment toward that individual and may act contrary to what strict principles of justice would normally warrant.

The second aspect of moral reasoning is that decisions are based on universal principles of morality. Kohlberg (see Crain, 1985) suggests that it is important to both protect individual rights and settle disputes democratically. This may be done successfully as individuals primarily observe a higher level of moral development that provides guiding principles for achieving justice. Crain (1985) further states that these principles require us to treat everyone the same; the principles are therefore universal.

Authoritative parenting tends to facilitate the development of moral reasoning more effectively than any other parenting style (Pratt et al., 1999). This development is likely because authoritative parents typically use induction and reasoning in disciplining their children, thus helping the children to internalize social values and norms (Grusec & Goodnow, 1994). As a child internalizes these values, he or she is more able to base decisions on these common values and principles, which is one of the two elements of moral reasoning. Empathy is also facilitated because as the parent explains his or her intentions and feelings about certain rules, the child has the opportunity to better understand another individual.

In contrast, authoritarian parenting tends to elicit fear, anger, and anxiety, and it is therefore associated with lower levels of moral development in children. The emotions elicited direct a child's attention toward external consequences and hinder the internalization of social values and the ability to have empathy for others. The permissive style of parenting does not actively hinder moral development, but neither does it provide children with sufficient opportunity to internalize values, which may impact moral development more indirectly (Lopez et al., 2001).

Hence, the relationship between parenting style and the degree of empathy—one aspect of moral reasoning—that children develop has been well established by previous research. It is has also been shown clearly that there is a link between empathy and distributive justice, particularly surrounding the allocation of punishment. However, the specific relationship between parenting style and punishment allocation has not been thoroughly investigated. Therefore, the present study will investigate this relationship. We believe that the degree of punishment allocated by participants will be influenced by the perceived parenting style of their parents.

Method

Participants
Participants were recruited from psychology courses at Brigham Young University (BYU) and received extra credit from professors who offered it.
There were a total of 84 volunteers: 49 females and 35 males. Most participants were of typical college age—generally ranging from about 20 to 30 years old. Participants were recruited through class visits and distribution of flyers.

**Materials**

*Informed Consent Form.* The informed consent form briefly introduced the study without disclosing its specific purposes, noted potential risks, assured confidentiality of participants, and provided contact information for participants with additional questions. The consent form was included as a coversheet to the packet of administered surveys.

*Household and Civic Measure of Punishment Allocation.* The Household and Civic Measure of Punishment Allocation (HCMPA) was developed for the present study. It included four different offense scenarios—two civic offenses and two household offenses. Specific details (e.g., race, gender, etc.) about characters in each scenario were not provided. The scenarios within each of the two categories differed only in the severity of the offense; all other elements, such as location, type of individuals involved, and reason for the dispute, were kept constant. A pilot study was conducted to ensure that the behaviors in the scenarios did, in fact, differ in their degree of severity. An example of one of the scenarios, the severe civic offense scenario, follows:

"An adolescent enters a gas station and gets into a dispute with the clerk over the amount of change received for a purchase. While they are arguing, the phone rings and the clerk turns away momentarily, leaving the cash drawer exposed. The adolescent strikes the clerk on the back of the head hard, leaving her unconscious and bleeding badly. He then empties the cash register and quickly exits the gas station."

The participant was then asked to indicate the degree of punishment that he or she felt was appropriate, using an anchored Likert scale ranging from 0 (no punishment) to 4 (maximum punishment). The participant was also asked to provide an example of the punishment that he or she felt would be appropriate. (To see the scale, please see Appendix A, HCMPA)

*Social Desirability Scale.* The Social Desirability Scale was developed by Crowne and Marlowe (1960) in order to assess the degree to which participants are prone to give socially desirable responses. It consists of 33 statements that participants may mark as either “True” or “False.” The questionnaire includes statements like the following: “Before voting I thoroughly investigate the qualifications of all the candidates.” This scale was included as a distracter; it was placed between the civic and household scenarios.

*Parental Authority Questionnaire.* The Parental Authority Questionnaire (PAQ) (Buri, 1991) contains 30 questions in reference to a respondent’s parent. Two versions of the PAQ were used in the present study—one for the respondent’s mother and another for the father. Aside from parent gender, the versions are identical in content and form. The 30 questions in each questionnaire include 10 questions measuring each of the three parenting styles. The following are examples of the statements used for each of the three parenting styles:

**Authoritative:** "As I was growing up, once family policy had been established, my father discussed the reasoning behind the policy with the children in the family."

**Authoritarian:** "Even if his children didn’t agree with him, my father felt that it was for our own good if we were forced to conform to what he thought was right."

**Permissive:** "While I was growing up, my father felt that in a well-run home the children should have their way in the family as often as the parents do."

A Likert scale ranging from 1 (strongly disagree) to 6 (strongly agree) is used by respondents to denote the level of accuracy the statement has for their parents. The test-retest reliability on the PAQ is .78 for mother’s authoritativeness, .86 for mother’s authoritarianism, and .81 for mother’s permissiveness. It is .92 for father’s authoritativeness,
.85 for father’s authoritarianism, and .77 for father’s permissiveness.

**Design and Procedure**

**Pilots.** In preparing to conduct research, two pilots were conducted to test the validity of the scenarios that were to be used in the experiment. Researchers wanted to ensure that the scenarios were dependably measuring the seriousness of the offenses being reviewed. For each pilot, the scenarios were adjusted because no significant difference in severity ratings was observed. For each pilot, the scenarios were rated on a Likert scale from 0 (requiring no punishment) to 6 (requiring maximum punishment). As with the final version of the survey, the participant was also asked to provide an example of the punishment that he or she felt would be appropriate. After each pilot was conducted, the results were examined and the scenarios underwent any necessary revisions to ensure that the scenarios were similar but differed effectively in the severity of the offenses being committed in each. The pilots were conducted with upper-level psychology students prior to conducting the primary research.

When the first pilot was conducted, the scenarios were passed out to 23 participants. They were asked to read the instructions and to complete the survey. The results of the pilot session were then briefly reviewed, and the participants were asked for verbal feedback on the scenarios; feedback was used in making alterations to the survey for subsequent pilots and for research. Though the household scenarios were left unchanged, the civic scenarios were adjusted because no significant difference between moderate and severe offenses was observed in the pilot’s results. Consequently, new versions of the civic scenarios were created, and a second pilot was conducted to evaluate the altered scenarios.

In the second pilot, 12 participants were given one version of the civic scenarios while 10 other participants were given a second version. This was done in order to identify which version would demonstrate a larger difference in severity ratings. The forms were subsequently collected from the participants and the results were reviewed. The second pilot demonstrated the following results: for the first altered version of the civic scenarios, $M = 3.58$ for the moderate civic scenario and $M = 6$ for the severe scenario. For the second altered version of the civic scenario, $M = 3.2$ for the moderate civic scenario; the severe civic scenario was similarly $M = 6$. Thus, the second version of the moderate scenario was chosen for the final version, as it demonstrated the greatest difference from the mean of the severe offense scenario. Once the pilots were completed and differing severity had been established for the different scenarios, the Likert scale was adjusted to range from 0 to 4.

**Study protocol.** The procedures implemented in conducting the research were the same in each of the research sessions. To begin each session, participants were provided with the opportunity to sign up for extra credit and were then given the research packets once everyone was ready to begin. Each research packet was divided into two sections. The first section contained the following materials in the listed order: informed consent form, the civic scenarios from the Household and Civic Measure of Punishment Allocation (HCMPA), the Social Desirability Scale (Crowne & Marlowe, 1960), and the household scenarios from the HCMPA. The second section of the research packet contained the Parental Authority Questionnaires (PAQ) for mothers and fathers (Buri, 1991).

**Instructions.** The research participants were next given the following verbal instructions: "Please read and sign the consent form, then detach and pass forward." Once the consent forms had been collected, everyone was given the following verbal instructions for the first part of the research packet:

"Please fill out the questions in Part 1 and read the instructions carefully. Stop at Part 2 for further instructions. Please do not talk out loud or interact with each other. If you need to ask a question, please raise your hand. When finished, please put your pencils down and look up."

After everyone had completed the first section, the participants were instructed as follows:
Results

Pre-Analysis

Civic and household scenario variables. It was decided that responses for both moderate and severe scenarios would be summed for each condition: civic and household. Combining participants' scores for the moderate and severe scenarios for each condition allowed us to look at the overall allocation of punishment. Each condition had one Likert-type response scale (0–4) for the moderate scenario and an identical response scale for the severe scenario. Responses to both scales were combined (with a new scale range of 0–8) and used for further analysis as one variable.

Parenting style variables. Preliminary analyses of the data from this study demonstrated that the authoritative parenting style was most powerful in influencing punishment allocations. After performing a tertile split on all three parenting styles, we found that the only significant effects involved authoritative parenting. Hence, we decided that variables would be constructed to contrast authoritative and non-authoritative parenting styles. The PAQ includes 30 items that demonstrate authoritative, authoritarian, or permissive parenting styles—10 items for each style. Only the authoritative subscale was used to differentiate between authoritative and non-authoritative parents. A score of 40 to 60 on this subscale classified a parent as authoritative, and a score of 10 to 30 classified a parent as non-authoritative. Scores between 30 and 40 were not used because they were neither clearly authoritative nor clearly non-authoritative.

Analysis

Section preface. Statistical tests were performed on data that involved both mothers and fathers. However, throughout the results, statistics concerning the parenting styles of mothers were not found to be statistically significant. Therefore, they are not discussed at length in the subsequent results.
**Discussion**

It was hypothesized that the degree of punishment allocated by participants would be influenced by the perceived parenting style used by the parents of the respondents. This is because parenting styles have been found to have an effect on the cognitive and social development of adolescents (Pratt et al., 1999). Our study demonstrated that if the father of the participant was authoritative, then the participant was more likely to allocate a higher level of punishment than if the father of the participant was non-authoritative.

With regard to the varying levels of punishment, it is interesting to note that the authoritative style of the participant's father was more predictive of punishment allocation in the household scenarios than in the civic scenarios. A potential hypothesis for this result is that the participants' views of government and law enforcement roles may have impacted the allocation of punishment in the civic scenarios more than parenting styles did. The qualitative recommendations of punishment provided by participants tend to support this idea.

In the household scenarios, participants attempted to designate more specific punishments, while in the civic scenarios, the recommendations generally relied on local law enforcement to actually choose and allocate a punishment.

Also, we noted that of the two household scenarios, the influence of the authoritative father on punishment allocation was only significant for the first, less severe scenario. It is possible that because the first household scenario was significantly less severe, the parenting style would have had a stronger impact on punishment allocation. Furthermore, it may be that the offense in the second scenario was so severe that extreme punishments would be allocated regardless of the parenting style.

Even though we found that fathers’ parenting styles significantly influenced punishment allocation by their adult children in many cases, we found no such effect for mothers’ parenting styles. It is not
clear why significance was found in the parenting styles of fathers but not of mothers. There are many possible hypotheses for this result. It may be that in many households, each parent punishes different offenses. The HCMPA may have only measured types of offenses that fathers were likely to punish, and therefore the mothers' influence was not significant. It may also be that many of the participants involved in this study generally perceived the father's influence to be greater than the mother's. There may also be some sort of interaction between fathers' and mothers' parenting styles that was not examined in this study.

It is additionally possible that children may look more to their fathers for guidelines on appropriate punishment allocation. Children often see their fathers as the enforcers of rules and the disciplinarians, while mothers are often viewed as the nurturing figures in the home (Goldman & Goldman, 1983). In our society men still predominantly make the rules, laws, and judgments—they may consequently be seen as "running" most of our society. If the father is the enforcer of the rules, then it is probable that children may use the father's methods for punishment allocation. This may explain why the mother's influence wasn't found to be statistically significant in this study. As the gender dynamic in our society is changing with time (see Lips, 2004), the gender effect may become less significant. Further changes to the HCMPA and different statistical tests may help in clarifying this difference between parents.

Finally, we hypothesized that of the two elements of moral development (empathy and decision-making based on universal principals or values), empathy would more strongly result from authoritative parenting. However, our results suggest otherwise. The results for authoritative fathers suggest that (1) we overestimated the impact of empathetic development, and (2) perhaps the internalization of social norms has a greater influence than we previously thought.

The children of authoritative parents are more likely to have internalized social norms and values than the children of non-authoritative parents (Gruszc & Goodnow, 1994). They are also more likely to develop empathy. However, from our sample it appears that social norms and values became more important than empathy. Strongly internalized values lead to an expectation that everyone will follow these values. When a child internalizes social norms and values, those norms and values become a part of that child's concept of how individuals ought to behave (Grusiec & Goodnow, 1994; Lopez et al., 2001). Thus, if a child saw an individual not behaving in accordance with his or her understood, internalized norms and values, the child would expect adverse consequences to follow. Hence, this child would be more likely to perceive the individual as deserving of punishment. This in turn may lead to the more severe punishments observed in the data.

Conclusions

We acknowledge that this study carries certain limitations. Firstly, the HCMPA contains only limited scenarios. It would, of course, be impossible to test for every possible offense in civil or household situations, but the HCMPA would certainly be improved if it were to contain a greater variety of scenarios. Our scenarios addressed only two specific kinds of scenarios. Additional scenarios could provide a broader spectrum of information that could better illustrate why, how, and under what circumstances parenting style influences punishment allocation.

Also, the nature of the population from which we drew our sample may have been an additional limitation. The majority of the students at BYU come from homes where The Church of Jesus Christ of Latter-day Saints is the predominant faith. A person is likely to find authoritative parenting among these parents because this type of parenting style is consistent with the religious beliefs of that faith. Because of this, our sample included a larger
number of participants from authoritative homes than from non-authoritative homes.

In addition to these limitations, it is possible that variables other than the perceived parenting style of the participants' parents influenced the severity of allocated punishment. As mentioned above, participants' views of law enforcement may have influenced the punishment allocation. Our study did not take into account other factors such as this that may influence responses.

Nonwithstanding these limitations, this study provides compelling results and invites further investigation. Additional research could include expanding the HCMPA to test for punishment allocation in a greater variety of situations. Further studies might also consider attitudes toward law enforcement, as well as other potential factors, that may influence a participant's response. A longitudinal study in which parenting styles were actually observed, rather than simply reported by adult children, would further establish and validate the relationship between parenting style and punishment allocation found in our study.

References


Appendix A

The Household and Civic Measure of Punishment Allocation (HCMPA)

Instructions: Read each of the following scenarios and circle the corresponding number, on the provided scale, that best describes the level of punishment you feel should be appropriately administered for the respective offense. Then provide an example of a punishment that you believe would be appropriate and equivalent to the score you provided. There are no right or wrong answers. We are looking for your overall impression regarding each situation.

Civic Scenarios

Scenario 1
An adolescent enters a gas station and gets into a dispute with the clerk over the amount of change received for a purchase. During the dispute, both parties are yelling. While they are arguing, the phone rings and the clerk turns away momentarily, leaving the disputed amount of money on the counter. The adolescent takes the money on the counter and quickly exits the gas station.

What degree of punishment do you believe would be appropriate in this situation?

No Punishment ______ Maximum Punishment

0 1 2 3 4

Please provide an example of a punishment that would be appropriate for this situation:

Scenario 2
An adolescent enters a gas station and gets into a dispute with the clerk over the amount of change received for a purchase. While they are arguing, the phone rings and the clerk turns away momentarily, leaving the cash drawer exposed. The adolescent strikes the clerk on the back of the head hard, leaving him unconscious and bleeding badly. He then empties the cash register and quickly exits the gas station.

What degree of punishment do you believe would be appropriate in this situation?

No Punishment ______ Maximum Punishment

0 1 2 3 4

Please provide an example of a punishment that would be appropriate for this situation:

Household Scenarios

Scenario 1
An adolescent and their parent are discussing an issue in their home. The conversation escalates into an argument and during the dispute both individuals are yelling. The adolescent becomes especially upset and yells, “I hate you!” and walks away.

What degree of punishment do you believe would be appropriate in this situation?

No Punishment ______ Maximum Punishment

0 1 2 3 4

Please provide an example of a punishment that would be appropriate for this situation:
Scenario 2
An adolescent and their parent are discussing an issue in their home. The conversation escalates into an argument and during the dispute both individuals are yelling. The adolescent becomes especially upset and yells, “I hate you!” then suddenly strikes the parent before walking away.

What degree of punishment do you believe would be appropriate in this situation?

No Punishment _______ Maximum Punishment

0 1 2 3 4

Please provide an example of a punishment that would be appropriate for this situation:
Retrograde amnesia refers to the memory loss of events occurring directly prior to a traumatic experience (Riccio, Millin, & Gisquet-Verrier, 1993). Numerous case studies have described people with signs of retrograde amnesia following concussive brain injury, seizure, encephalitis, stroke, aneurysm, and chronic alcohol abuse (Riccio, Millin, & Gisquet-Verrier, 1993). Viewing shocking photographs (Schmidt, 2002) and experiencing stressful situations such as skydiving have also proven to cause retrograde amnesia (Thompson, Williams, L’Esperance, & Cornelius, 2001). However, few studies have investigated how witnessing a traumatic event (as opposed to participating in the event) may cause retrograde amnesia. One study has shown a possible connection, indicating that when distractions are presented simultaneously with target information, participants have difficulty recalling or identifying the target information (Levy, 1998).

When an event is manipulated in such a way as to make it more dramatic or emotionally charged, it will leave a stronger impression on the subject. Thus, the subject will more accurately recall that particular event, but the recall of events occurring immediately beforehand can be affected negatively (Berntsen, 2002).

One experiment that has given some evidence that retrograde amnesia can occur after a person has witnessed a traumatic event was conducted by Loftus and Burns (1982). This experiment involved showing two independent groups of participants two different versions of a filmed robbery, one where a young boy was shot in the face and one in which the robbers left without shooting the boy. The participants who viewed the more emotionally charged version (with the shooting of the boy) had a stronger and more accurate recollection of the scene; however, they were less able to recall details that were shown directly before the boy was shot—thus demonstrating symptoms of retrograde amnesia. Specifically, 27.9% of the participants watching the nonviolent version of the film recalled the number on the boy’s jersey, whereas only 4.3% of the participants watching the violent version of the film were able to recall the number on his jersey (Loftus & Burns, 1982).

Another study (Ihlebaek, Love, Eilertsen, & Mangnussen, 2003) investigated the effects of...
witnessing a traumatic event in vivo versus the effects of witnessing a traumatic event via video presentation. This study showed that witnessing real-life events may be more traumatic than witnessing events on film. It also showed that witnesses viewing videos remember more details than witnesses viewing real-life situations. Furthermore, these findings suggest that participation in traumatic events may cause retrograde amnestic effects.

The aforementioned studies were designed to study the relationship between retrograde amnestic symptoms and witnessing severe traumatic events. Further research is necessary, however, in order to validate Loftus and Burns's (1982) findings for mildly traumatic events. Thus, this experiment was designed to investigate the relationship between witnessing mildly traumatic events and retrograde amnestic symptoms. This relationship was explored by assessing recall of names presented to participants prior to a mild traumatic event—a researcher falling and being injured in the presence of the participants. We predicted that participants in the experimental (trauma) group would be less able to recall the previously presented names than participants in the control (non-trauma) group.

Method

Participants

There were 85 undergraduate student participants from psychology classes at Brigham Young University in our study, 53 in a control group (the non-trauma group) and 32 in an experimental group (the trauma group). Participants took part in the study on a volunteer basis and some were offered extra credit at their instructor's discretion. All participants were informed that by taking the questionnaire they were consenting to participate in our study.

Materials

To parallel the Loftus and Burns (1982) study, we created a similar questionnaire that was applicable to this experiment. The questionnaire consisted of 10 questions that required participants to recall recent events in their lives. Similar to the Loftus and Burns study, filler questions were included in our questionnaire to make it less obvious to the participants what we were studying. The eight filler questions used in our study asked about things such as what the participants had eaten for breakfast and when their last date was. The two questions of interest asked participants to recall the presented researchers' names, Cassie and Sarah. The first question asked the name of the researcher who was the subject of the traumatic event for the experimental group, Sarah. The second question asked the name of the researcher who gave the introduction, Cassie. The results from this study were based upon the participants' responses to these two questions of interest—that is, we operationalized retrograde amnestic symptoms as occurring when participants were unable to correctly recall these two names. This is similar to the Loftus and Burns study, wherein the question of interest asked participants to recall the number on the child's jersey.

Design and Procedure

The questionnaires were given in four separate classes over a two-week period, with two classes being designated the control group and the other two the experimental group. In each condition the researchers wore the same attire, took the same amount of time to speak and introduce themselves and the study, and gave the same instructions. During the instructions, the researcher described the questionnaire and told participants how to appropriately answer the questions. Following the instructions, the researchers introduced themselves.

In the experimental condition, participants witnessed the researcher, Sarah, fall down the stairs after introducing herself, violently striking her face on the ground. They then observed Sarah getting up from the fall, appearing to bleed while shaking from shock. Both Cassie's and Sarah's names were said during the presentations. In the experimental group, Cassie's name was said approximately 7
seconds before the fall, and Sarah's name was said 1 second before the fall. The control group was given the same presentation without the mildly traumatic event—the researcher falling.

After the presentation, the questionnaires were administered while the two researchers (Cassie and Sarah) exited the room. (In the experimental condition, Cassie escorted Sarah to the restroom.) Each of the participants were given as much time as needed to complete the questionnaire. After the last questionnaire was returned, the researchers re-entered the testing room. The researchers then debriefed the participants, explaining the true nature of the study, followed by a brief question and answer session. In the experimental condition, the participants were informed that the accident was staged. The entire procedure took approximately 10 to 15 minutes.

Results

First, the data from the two questions of interest were converted into percentages (see Figure 1). Four chi-square analyses were then performed on the differences in percentages. The first chi-square test compared the percentage of participants who remembered Sarah’s name in the control group versus those that correctly recalled her name in the experimental group. The second test compared the percentage of participants who correctly recalled Cassie’s name in the control group versus those who correctly recalled Sarah’s name in the experimental group. The third test compared the percentage of participants who correctly recalled Cassie’s name versus those who correctly recalled Sarah’s name in the experimental group. The fourth test compared, within the control group, the percentage of participants who correctly recalled Cassie’s name versus those who correctly recalled Sarah’s name (see Figure 2).

Results from the four chi-square tests show strong evidence for a relationship between witnessing a traumatic event and experiencing retrograde amnestic symptoms. The first test, which compared the percentages of participants in the two different groups who correctly recalled Sarah’s name, had the greatest statistical significance. In the control group, 66.6% of participants were able to recall Sarah’s name, which is approximately the same percent who remembered Cassie’s name. In contrast, only 21.8% of the participants in the experimental group were able to recall Sarah’s name. The difference in percentages was significant; \( \chi^2(1, N = 85) = 40.8, p < 0.001 \).

The second chi-square test, comparing the percentages of participants in both groups who correctly recalled Cassie’s name, also indicated that witnessing a mildly traumatic event may be related to retrograde amnestic symptoms. Of the participants in the control group, 69% correctly recalled Cassie’s name, whereas in the experimental group, only 50% of the participants correctly recalled her name. There was a significant difference between the percentage of participants who recalled Cassie’s name in the control group versus the experimental group, just as was seen with Sarah’s name, \( \chi^2(1, N = 85) = 7.5, p < 0.01 \).

The third chi-square test compared the difference between the percentages of participants who remembered Cassie’s name versus Sarah’s name in the experimental group. In the experimental group, 50% of participants remembered Cassie’s name, whereas only 21.8% remembered Sarah’s name.
This chi-square test indicated a significant difference between Sarah's name being correctly recalled less often than Cassie's name, \( \chi^2(1, N = 32) = 17.3, p < 0.001 \). This finding suggests that the closer to the traumatic event the detail (name) is presented, the more likely the witness may be unable to recall that detail.

The last chi-square test measured the difference between the percentages of participants who correctly recalled Cassie's name versus Sarah's name in the control group. In the control group, 69% remembered Cassie's name and 66.6% remembered Sarah's name. The chi-square test indicated that there was no significant difference between the two percentages, \( \chi^2(1, N = 53) = 0.13, p > .1 \) (see Figure 2).

**Discussion**

Our prediction, that participants would be less likely to recall the names of researchers after witnessing a traumatic event, was supported by our results. A significant difference was identified in three of the four analyses. Using chi-square tests, we found that Sarah's name and Cassie's name were both correctly recalled less in the experimental condition than in the control condition. A significant difference was also found between the number of participants who correctly recalled Sarah's and Cassie's names in the experimental condition. No significant difference was found, however, between correct recall of Sarah's and Cassie's names in the control condition—participants remembered both names equally well. Statistically, this study indicates that there is a relationship between witnessing a mildly traumatic event and having retrograde amnestic symptoms. Furthermore, this study demonstrates that the closer the target information item is to the traumatic event, the more likely it is to be incorrectly recalled.

Possible confounds of this study include the ambiguity of the questionnaire and the number of participants observing the event in each administration of the study. The participants were not asked if the questionnaire was confusing or difficult to understand; however, we suggest that the questions may not have identified the researchers clearly enough, so that the participants would know to whom the question was referring. Because of this, participants may have been confused as to which question was referring to Cassie and which was referring to Sarah. Thus, modifying the questionnaire to eliminate ambiguity may increase the validity of the experiment.

Another possible confound was the number of participants in each administration of the study. The experiment was administered in four psychology courses; however, it was not administered to the same number of participants each time. Thus, some participants may have experienced something similar to the bystander apathy effect, which can occur when a person is slower to provide help to an individual in distress when there are other bystanders.

### Figure 2

<table>
<thead>
<tr>
<th>Relationships</th>
<th>Percentages</th>
<th>( \chi^2 ) Value</th>
<th>Significance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sarah (Control) vs. Sarah (Experimental)</td>
<td>66.6% vs. 21.8%</td>
<td>40.68</td>
<td>( p &lt; 0.001^{***} )</td>
</tr>
<tr>
<td>Cassie (Control) vs. Cassie (Experimental)</td>
<td>69% vs. 50%</td>
<td>7.5</td>
<td>( p &lt; 0.01^{**} )</td>
</tr>
<tr>
<td>Cassie (Experimental) vs. Sarah (Experimental)</td>
<td>50% vs. 21.8%</td>
<td>17.28</td>
<td>( p &lt; 0.001^{***} )</td>
</tr>
<tr>
<td>Cassie (Control) vs. Sarah (Control)</td>
<td>69% vs. 66.6%</td>
<td>0.13</td>
<td>( p &gt; 0.10 )</td>
</tr>
</tbody>
</table>
(Garcia, Weaver, Moskowitz, & Darley, 2002). It is possible that the traumatic event may have been more or less traumatic depending on the number of bystanders (participants). We suggest according to this theory, that groups with fewer participants most likely experienced the event as more traumatic than groups with more participants.

The primary limitation of this study was the lack of participant diversity. The participants were recruited via convenience sampling and thus the sample may have been misrepresentative of all BYU undergraduates. This limitation may have been corrected for by using probability sampling procedures.

Furthermore, though the findings of this experiment tend to support the findings of previous research, they also suggest that other associated variables and possible explanations should be investigated in future studies. For example, future research might fruitfully examine other possible explanations for the data such as whether the inability to recall information presented before the traumatic event was due to an actual forgetting (amnesia) or some other factor such as inability to decode the information into long-term memory due to the distraction of the event. In addition, future research could investigate why information presented immediately prior to traumatic events is less likely to be correctly recalled than information presented earlier. Studies have shown that in some cases, traumatic events have created a tunnel effect for witnesses, where they report central details of the event rather than peripheral details (Berntsen, 2002). This could have an impact in participant recall of details prior to or surrounding an event compared to details of the event itself. Additionally, future research should investigate the length of time that an individual’s recall is impaired by retrograde amnestic symptoms, something that this study failed to consider. Because Cassie’s name was said only seven seconds prior to the traumatic event in the experimental group, it is evident that the retrograde amnesia effect from witnessing a traumatic event can not only happen directly prior to the event (approximately 1 to 3 seconds), as seen with Sarah’s name, but may also extend 7 to 10 seconds prior to the event. The data seem to support the conclusion that the likelihood of being unable to recall details increases as the details are given closer to a traumatic event.

Future research should also continue to investigate the effects of different qualitative levels of traumatic events (i.e., mild or severe) on memory recall. Increasing or decreasing the intensity of the traumatic event may have an effect on the length of time and intensity of the retrograde amnestic symptoms (Dutton & Carroll, 2001). Though we considered our event less traumatic than Loftus and Burns’s (1982) event, studies have shown that often what participants think is traumatic differs from what the researchers may label as traumatic (Bohanek, Fivush, & Walker, 2005); thus, future research should specifically explore the relationship between participants’ perceptions of traumatic events and the length and intensity of their symptoms.

Although there is much evidence to support our hypothesis, we suggest that it would be beneficial to increase the understanding of retrograde amnestic symptoms by investigating the aforementioned associated variables. Thus, though the findings of this study support previous findings on retrograde amnesia, such as those of Loftus and Burns (1982), they also call for further investigation into previously unstudied areas.

References


Evidence-Based Practice
Movements in Psychology:
Empirically Supported Treatments,
Common Factors, and Objective
Methodological Pluralism

Evidence-based practice is arguably the most consequential and controversial movement in psychology today (Norcross, Beutler, & Levant, 2005). Although everyone agrees that practice should be informed by evidence (Westen & Bradley, 2005; Norcross et al., 2005), there is much disagreement about what qualifies as evidence (e.g., Reed, 2005; Kihlstrom, 2005; Messer, 2005) and whether certain evidence-based protocols extend to real-world practice (e.g., Westen, 2005a; Stiman & DeRubeis, 2005). This conflict is not a simple dispute to be resolved in the laboratory—it is a "culture war" between different worldviews in the quest for truth, respectability, and economic inclusion (Norcross et al., 2005, pp. 7–8).

Considering its controversial and consequential nature, the evidence-based "war" carries both promises and pitfalls. On the one hand, such evidence-based considerations remind researchers and practitioners of the need to be accountable for their research methods and types of practices. This reminder ought to be refreshing, considering psychotherapy's history of unexamined and uninvestigated methods and theories (Slife, Wiggins, & Graham, 2005). As scientific researchers and practitioners, we ought to be open to the critical examination of all types of methods and practices, however prized or popular they may be. Moreover, a commitment to critical examination hinders psychology from becoming a relativistic discipline where anything goes (Slife, Wiggins, & Graham).

On the other hand, pressure for evidential basis could potentially lead to an evidence-based framework that uncritically rests upon a limited conception of evidence. Considering the intense disagreement in psychology about what qualifies as evidence, it
is plausible that an evidence-based framework could be biased toward the most popular, powerful, or lucrative conception. This preliminary bias might hinder an examination of the conception itself, such as its philosophical assumptions and whether its methods are best suited for the needs of actual practice. Worse, it would likely subordinate or disenfranchise alternative conceptions of evidence and the methods and practices they imply—because such do not conform to the criteria of the chosen framework, they would be ruled out before an investigation even begins. For these reasons, a framework for evidence-based practice must itself be validated on scientific grounds and should be wary about committing to an overly narrow conception of evidence.

The challenge for psychology, then, is to articulate an evidence-based framework that is both inclusive and objective. It should be sufficiently open and flexible to include the utilization and creation of all necessary research methods and types of practice. At the same time, its desire for inclusion cannot slip into an anything-goes relativism—a commitment to evidential basis requires objective examination of research methods and practices, according to scientific standards.

Unfortunately, such a framework does not currently operate in psychology. Although two prominent evidence-based movements have emerged in the past decade, neither allows for a framework that is genuinely objective and inclusive. This article discusses both movements—the empirically supported treatment (EST) movement and the common-factors movement—contending that neither provides a framework that achieves both objectivity and diversity. To the contrary, each is based on an unexamined conception of evidence that commits a preinvestigatory bias against certain types of methods and practices—that is, certain methods and practices are ruled out before an examination even begins. After introducing each framework and briefly demonstrating each’s respective limitations, I will contend that a third framework—objective methodological pluralism—is better suited at offering an objective and inclusive framework for evidence-based practice.1

### Empirically Supported Treatment (EST) Movement

The empirically supported treatment movement (EST) defines evidence-based practice as the application of specific, often manualized, treatments that comply with stringent standards of experimental verification (Butcher, Mineka, & Hooley, 2004; Norcross et al., 2005; Safran, 2001). Generally, these treatments are tested according to their efficacy in treating specific psychological disorders, as identified in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM; Westen, 2005b).

The EST movement sprang from increasing pressure to justify psychological practices with evidence. Although psychology has always been concerned with evidence-based practice (Norcross et al., 2005), recent years have brought increasing pressure for keeping pace with medical treatments such as prescription drugs (APA, 2006). In an age of increasing anxiety for the justification of health care practices (Norcross et al.), psychological treatment has developed a reputation—however unjustified—of being less reliable and substantiated than medical treatment (APA; Westen & Bradley, 2005).

In an effort to remedy psychology’s second-class citizenship, a succession of task forces within Division 12 (Clinical Psychology) of the American Psychological Association (APA) sought to bolster the scientific validity of specific psychological treatments (APA, 2006). Beginning in 1993, the

1. For more extensive information, see Wendt, D. C. (2006). Evidence-based practice movements in psychology: Empirically supported treatments, common factors, and methodological pluralism. [Honors thesis.] Available at BYU’s Harold B. Lee Library and Department of Psychology (1001 SW/KT).
division “constructed and elaborated a list of empirically supported, manualized psychological interventions for specific disorders” (Norcross et al., 2005). Empirical support for these interventions—commonly referred to as empirically supported treatments (ESTs)—is demonstrated through rigorous laboratory testing of a large random sample, in which each participant is randomly assigned to an experimental group (treatment) or a control group (no treatment, or “wait-list”). By controlling for all variables other than the treatment in question—through random assignment and manualized protocols—it is thought that successful treatments can be experimentally isolated and identified (Butcher et al., 2004).

This experimental method is called a randomized clinical (or controlled) trial (RCT), widely esteemed as the gold standard of research evidence in medicine (Norcross et al., 2005; Safran, 2001). The RCT’s prevalent use in medicine is not surprising, considering that the RCT is tailored to fit traditional medicine’s theory of disease and treatment: disease is the presence of one or more specific symptoms, and treatment is a specific, uniform procedure for alleviating such symptoms (Bohart, O’Hara, & Leitner, 1998). RCTs are especially esteemed for determining prescription drug efficacy, in which an experimental group is given the treatment drug and a control group is given a placebo. If symptom-alleviation is significantly higher in the experimental group than in the placebo group, then it is thought to be the treatment—not the patient, the doctor, a placebo effect or happy chance—that is responsible for the change. Third-party payment providers can then trust the treatment to be universally effective and subsequently offer coverage for such.

Reportedly, Division 12’s intention was merely to establish that psychological treatments can be as or more effective than medical treatments, not necessarily to produce a monopolistic list of acceptable treatments (APA, 2006). Nonetheless, the division’s project sparked considerable interest in researching and implementing specific treatments, with the hope, in some circles, that psychological practice could be grounded upon a framework consisting solely of ESTs (APA). In fact, some professionals in psychology began contending that “empirically supported treatments are all the profession should allow patients to choose” (Bohart, 2003, p. 1), and many payment providers and funding agencies began limiting funding for certain disorders to ESTs (APA; Norcross et al., 2005; Safran, 2001).

As a result, the EST movement has become so entrenched in psychology that “evidence-based practice” has grown in some circles, to be synonymous with “empirically supported treatments” (Westen & Bradley, 2005, p. 266). With this equation in place, ESTs and RCTs are not seen as a way to justify psychological methods and practices, but the way.

By conforming to the strict criteria required by RCTs, an EST framework succeeds at establishing a certain level of reliability and credibility. It ultimately fails, however, to be an objective and inclusive framework for evidence-based practice because it is uncritically biased toward a medical-model of treatment that is inconsistent with the needs of many real-world patients and circumstances (Westen & Bradley, 2005; Messer, 2004; Bohart et al., 1998). By being committed, a priori, to a medical-model of treatment, an EST framework (a) demands, without rationale, that real-world practice be reshaped to fit the logic of the RCT and (b) rejects alternative conceptions of evidence and the methods and practices they imply, not because of their potential fruitfulness, but merely because they do not conform to RCT criteria. This section will address these two factors in turn.

Shaping Practice to Fit RCT Criteria

By assuming that RCTs are the only appropriate method for evidence-based practice, an EST framework demands that clinical practice be shaped to fit methodological demands, as opposed to the other way around. In other words, an EST-framework is driven by method, regardless of the real-world situation. This is important, considering the fact that ESTs lack external validity to the majority of real-world patients and situations (Westen & Bradley,
2005; Bohart, 2003). Unlike the real world, RCT patients are limited to those whose symptoms are textbook cases of a single DSM disorder—thus, their results “may apply only to a narrow and homogeneous group of patients” (Butcher et al., 2004, p. 563).

From a mainstream experimental standpoint, the exclusive use of RCTs is understandable—one must eliminate confounding variables such as the presence of other disorders—but from a clinical perspective, the majority of real-world patients cannot be pigeonholed into a single diagnostic category. In fact, the majority of U.S. patients are comorbid (Morrison, Bradley, & Westen, 2003; Westen & Bradley, 2005), meaning they are diagnosed with two or more DSM disorders (Butcher et al., 2004). According to the National Comorbidity Survey (NCS), 56% of patients have had three or more disorders (Butcher et al.).

In response to this question of external validity, advocates for an EST framework insist, without evidence, that practice can be shaped for comorbid patients in a way that is consistent with RCT methodology (Morrison et al., 2003; Bohart, 2003). To make this claim, an EST framework takes its cues from medicine’s common practice of prescribing multiple drugs: the answer lies in treating each disorder on an individual basis. Just as a patient with multiple medical problems can take multiple kinds of prescription drugs, so can a patient with comorbid mental health disorders receive multiple treatments. A person who is diagnosed with both depression and anxiety, for example, would receive two manualized ESTs, one for each disorder (Morrison et al., 2003).

At face value this practice may seem plausible, but it rests upon a presumptive assumption of the medical-model, at least as far as it relates to psychology. This assumption is the atomistic assumption of comorbidity (my term), in which disease can be operationalized as one or more self-contained disorders. This allows for comorbid disorders to be understood, diagnosed, and treated on an individual basis. In other words, comorbid patients can be treated as if they have only one disorder (Morrison et al., 2003). Thus, one patient’s combination of depression and anxiety is not considered to be a qualitatively unique whole, but merely the quantitative sum of two self-contained parts (Morrison et al.). With this assumption, researchers and practitioners can presume that a single disorder for a comorbid patient manifests itself in the same way as it would for a noncomorbid patient—thus, both patients can be treated with the same RCT-verified, manualized treatment (Morrison et al.).

The atomistic assumption is method-driven because it forces real-world practice (treating a comorbid patient) to fit the logic of the medical-model (via the atomistic assumption), without considering whether this requirement is justified by evidence. If the atomistic assumption were to be evaluated, it would need to be from a wider conception of evidence than mere RCTs: because RCTs do not include comorbid patients, their results alone do not determine whether ESTs are appropriate for comorbid patients. Thus, critics of an EST framework have appealed to a broader framework of evidence—one that can more aptly inform how RCTs relate to actual practice—to show the problems of the atomistic assumption. This framework includes empirical research that suggests that ESTs have limited success for comorbid patients (Morrison et al., 2003; Messer, 2004).

The revelation of the atomistic assumption of comorbidity is just one way that critics have exposed the external validity limitations of an EST framework. Considerable research has demonstrated, for example, that manualized ESTs often hinder the presence of important therapeutic factors, such as the therapist’s genuineness, creativity, and motivation, as well as the patient’s faith in the therapist and the strength of the therapeutic relationship (Piper & Ogrodniczuk, 1999). Other research has shown that success rates of the RCT are inflated due to its disregard of long-term relapse rates and its exclusion of early dropouts before determining efficacy rates (Messer, 2004).
The preceding research casts doubt on the objectivity of an EST-framework, suggesting that such is based on an inherently limited and biased conception of evidence, as opposed to an infallible window to reality. This could be known, however, only by relying upon a conception of evidence that includes, but is not limited to, the RCT.

Rejecting Alternative Conceptions of Evidence

In connection with reshaping real-world practice to match the criteria of a medical-model of evidence, an EST framework fails to consider alternative conceptions of evidence and the methods and practices they imply, not because of their potential fruitfulness, but merely because they do not conform to the medical model’s presumptions of disorder, treatment, and human change. Such types of practice include, but are not limited to, humanistic therapies (client-centered, existential, experiential, gestalt), psychodynamic therapies (Freudian psychoanalysis, interpersonal, self psychology, object relations), and certain marital/family therapies (family systems, structural family; Butcher et al., 2004). Although these therapies are considerably different, they are similar in that they are not concerned, or solely concerned, with treating specific disorders using manualized treatments.

From the perspective of these other therapies, evidence-based practice would be significantly broader than a medical-model approach. Humanistic therapies, for example, are concerned with nondiagnostic issues, such as expanding a patient’s “awareness” and dealing with problems of “alienation, depersonalization, loneliness, and a failure to find meaning and genuine fulfillment” (Butcher et al., 2004, p. 584). A key component of humanistic research and practice is the relationship between the therapist and the patient. As mentioned above, a reliance on manualized treatments impedes the development of a therapeutic relationship, which is essential for humanistic therapies. To cultivate a therapeutic relationship, therapists must remain free to incorporate their own clinical wisdom, according to the specific patient and context. Because of the uniqueness of each patient and therapy session, good practitioners do not robotically apply step-by-step protocols, but instead “are artists who learn how to apply [empirically supported] principles in creative ways” (Beutler, as qtd. in Bohart, 2003, p. 4).

In contrast, a medical-model approach puts an emphasis on universal, “packaged” treatments for well-defined, compartmentalized disorders (Bohart et al., 1998). As mentioned earlier, this model takes its cues from medicine’s study of pharmaceutics, where “one must specify the treatment and make sure it is being applied correctly” (Bohart et al., p. 143). Just as a drug prescription is a specific, portable package in terms of its encapsulated ingredients and usage directions, so must a psychological treatment be “packaged” as an instruction manual with specific procedures and directions. In both cases, every patient receives the exact same thing, and it is this thing that is the agent of change—the health professional is merely a delivery person, and the patient, a passive recipient.

As long as an EST framework is built upon the assumption that packaged treatments—not therapists or patients—are responsible for change, its research agenda will commit a preinvestigatory bias against humanistic and other therapies; such therapies would be ruled out before investigation even begins. The disenfranchisement of these therapies would be troubling for many, if not most, psychotherapists, considering the discipline’s widespread eclecticism and integrationism (Slife & Reber, 2001).

Common-factors Movement

Because of the biases and exclusiveness of an EST framework, many researchers and practitioners have vehemently opposed restricting evidence-based practice to ESTs (e.g., Bohart, 2003; Greenberg & Watson, 2005; Messer, 2004; Morrison et al., 2003; Westen & Bradley, 2005).
Common-factors advocates\(^2\) have argued that a focus on specific treatments for specific disorders is only one way of conceptualizing psychological practice (e.g., Bohart, 2003; Westen & Bradley, 2005; Messer, 2004; Safran, 2001). An alternative approach is to discover and validate factors of therapeutic change that are common across treatments. These “common factors” include therapist techniques and characteristics such as empathy (Bohart et al., 1998), patient characteristics such as “active self-healing abilities” (Bohart, 2005, p. 218), and the relationship between therapist and patient (Norcross & Lambert, 2005; APA, 2006).

An attention to common factors is based on a different worldview than an EST framework, in which responsibility for change is not attributed to a specific treatment alone, but also to the therapist, the patient, and their dynamic relationship (APA, 2006). Psychotherapy researchers have claimed to measure and empirically validate common factors, and argue that the presence of these factors is often a more successful determinant of change than is mere adherence to ESTs (Norcross et al., 2005). Thus, from this view, the answer for evidence-based practice is to discover and apply common factors within all types of therapy, not impose a one-size-fits-all strategy (Bohart, 2003; Westen & Bradley, 2005).

A common-factors approach is appealing, considering that the majority of practitioners consider themselves eclectics and integrationists who value a wide array of research methods, therapeutic techniques, and theoretical orientations (Slife & Reber, 2001). Eclectic and integrationist practitioners believe that openness to a pluralism of methods, including ESTs, allows them to better help the wide range of patients and problems they encounter (Slife & Reber).

Because of its wide appeal, the common-factors movement has made a considerable impact in recent years. Most recently, the common-factors movement has played a significant role in shaping APA’s new policy on evidence-based practice in psychology (EBPP). The policy’s supplementary report, authored by the APA Presidential Task Force on Evidence-Based Practice (2006), makes it clear that a comprehensive strategy for EBPP “will consider [many] determinants [of effective practice] and their optimal combinations,” such as “the treatment method, the individual psychologist, the treatment relationship, and the patient” (APA, 2006, p. 275). The report’s endorsement of these and other common factors shows that this APA policy is in line with the worldview of the common-factors movement: responsibility for change is not attributed to a specific treatment alone, but also to the therapist, the patient, and their dynamic relationship.

Consistent with the common-factors movement, the APA policy explicitly values the inclusion of a diversity of methods. By “start[ing] with the patient” (APA, 2006, p. 273), the APA policy is thought to free itself from relying on a single method’s worldviews about the nature of illness and effective treatment. This allows it to be informed from a diversity of methods, including—but not limited to—ESTs (APA) and their respective conceptions of evidence. This accommodation to diversity is not surprising, considering APA’s goal for the policy to include a consideration of each valid perspective in the discipline (“A presidential,” 2005).

In comparison to an EST framework, the common-factors movement (including the APA policy) is based on a wider conception of evidence that allows it to more objectively avoid bias and better match the diverse needs of the discipline. After a close evaluation, however, it is clear that the APA policy is also a biased framework and is

---

2. These advocates represent a pluralism of views, each with a variety of nuances (Messer, 2004), but they are similar in their concern about an EST-framework’s “emphasis on specific treatment effects as opposed to common factors that account for much of the variance in outcomes across disorders” (APA, 2006, p. 272). Due to this similarity, I have conceptualized this group as representing a single movement—the common-factors movement.
not diverse enough for the wide range of practice in psychology. Just as an EST framework uncritically restricts acceptable evidence to a single method (RCT), so does the APA policy uncritically restrict acceptable evidence to a single epistemology.

The epistemology on which the policy is based is a narrow brand of empiricism. According to this epistemology, "we can only know, or know best, those aspects of our experience that are sensory" (Slife, 2006; see also Slife et al., 2005, p. 84). Given popular demands that natural science disciplines be grounded in empiricism, the policy's exclusion to empirical methods is understandable. As I will show, however, this exclusion is not based on evidence—analagous to a desire for RCTs, this APA policy merely assumes that empiricism is the only appropriate epistemology for evidence-based practice, in spite of the existence of other promising epistemologies. This mistake is consistent with much of psychology's history (Slife et al.), in which empiricism has been misunderstood as meaning objective or impartial, "in the sense of exposing what is actual or real" (Slife et al.). In other words, empiricism has not been seen as a particular epistemology or philosophy at all, but as a transparent window to the way things are (Slife et al.).

The APA policy perpetuates an equation of empiricism with reality, seeing no need to provide a rationale for its repeated, implicit equation of "evidence" with "empirically supported." The report document (APA, 2006) claims, for example, that "the purpose of EBPP is to promote effective psychological practice . . . by applying empirically supported principles of psychological assessment,

3. This conception of empiricism is a fairly traditional one and is the way the term is typically used in psychology. More liberal usages of empiricism differ somewhat, such as William James's radical empiricism, which encompasses "the whole of experience," including non-sensory experiences such as thoughts, emotions, and spiritual experiences (Slife, 2006). In mainstream psychology, however, the term empiricism is commonly used to refer to sensory experience only. Thus, throughout this paper, I will use empiricism to refer to sensory experience and non-empiricism to refer to non-sensory kinds of experiences.

case formulation, therapeutic relationship, and intervention" (APA). Here, as in several other places, the task force asserts that it endorses the application of empirically supported principles, but fails to explain why. In fact, nowhere in the policy or in its underlying report is a rationale given for a commitment to empirical research, and nowhere is a consideration given for even the possibility of the contribution of "non-empirical" research to EBPP. If the task force does in fact view empiricism as a particular epistemology, nowhere does it justify, or even explicate, its exclusive commitment to it.

This is a curious omission. If EBPP is exclusively committed to a single epistemology, why not come right out and say it? Indeed, why not call the movement empirically based practice in psychology? Perhaps the task force wants to have its cake and eat it too—to cater wholeheartedly and uncritically to one epistemology (empiricism) but talk about it in a way that implies it does not see it as an epistemology at all. This is similar to an EST framework and its equation of empirical support as RCT-verified. The problem with both frameworks is their equation of a narrow conception of evidence with a broader reality. An EST framework is exclusively committed to RCTs but assumes that such encompasses the broader field of empirical support (hence the broad designation, empirically supported treatments). Similarly, the APA policy is exclusively committed to empiricism but assumes that such encompasses the broader field of evidentiary basis in psychology (hence the broad designation, evidence-based practice in psychology).

Thus, the problems with the APA policy's commitment to empiricism are comparable to those from an EST framework's commitment to the RCT. In particular, the APA policy (a) demands, without evidence, that clinical practice be shaped to fit an empiricist epistemology and (b) marginalizes the inclusion of non-empirical epistemologies and their methods and practices, not because of their potential fruitfulness for evidence-based practice, but merely because they do not conform to the logic and criteria of empiricism.
Shaping Practice to Fit Empiricist Criteria

By assuming that empiricism is the only appropriate epistemology for evidence-based practice, the APA policy demands that clinical practice be shaped to fit empirical criteria. However, just as the majority of real-world patients cannot be neatly categorized as a number of disorders, so are the majority of real-world experiences not able to be neatly categorized according to the logic of empiricism. In fact, the vast majority of real-world phenomena of interest for evidence-based practice are not, strictly speaking, empirical in nature; rather, they are unobservable experiences, meaning they are not sensory in nature (Slife et al., 2005). Such phenomena include many, if not all, of the common factors of therapeutic change, especially those concerning the therapeutic relationship (Slife et al.).

Common-factors advocates and the APA task force have endorsed the investigation and implementation of such unobservable experiences for evidence-based practice (APA, 2006), but only in a way that is consistent with the logic of empiricism. From an empiricist standpoint, the way to handle an unobservable experience is to “operationalize” it. An operationalization is an observable, quantitative set of operations intended to represent an unobservable construct. For example, one might operationalize depression as a certain score on a questionnaire, or intelligence as a score on an intelligence test. Although operationalization is widely considered to be essential for the reliability and progress of a scientific discipline, this claim is not at all based on empirical evidence. Indeed, there is no empirical way of knowing whether, or in what manner, an operationalization relates to the original, unobservable, construct of study.

The APA policy ignores this problem, however, asserting that “good practice and science call for the timely testing of psychological practices in a way that adequately operationalizes them using appropriate scientific methodology” (APA, 2006, p. 274). No justification is given for this claim—it is merely assumed a priori. In this respect, the uncritical demand for operationalization is as problematic as the insistence that ESTs work for comorbid patients. Likewise, it rests upon an assumption that is as unfounded as the atomistic assumption of comorbidity.

This assumption is the observability assumption of unobservables (my term)—that unobservable meanings and phenomena can in fact be operationalized into observable phenomena (Slife et al., 2005). Like the atomistic assumption, the observability assumption is a preinvestigatory bias that does not sufficiently represent real-world phenomena. Similarly, there appears to be no justification for it at all, apart from its conformity to the logic of empiricism. Indeed, one cannot resort to empirical evidence on this matter, because it is not an empirical question—this would be akin to justifying the atomistic assumption with RCT evidence. If one were to justify the observable assumption, it would require theoretical expertise concerning the nature of unobservable meanings—and just as the RCT is not concerned with comorbid patients, so is empiricism not concerned with unobservables.

If one assumes, for example, that an operationalization is connected or related in any way to an unobservable meaning, this connection must itself be a non-empirical, unobservable one. Suppose, for example, a researcher seeks to study “happiness” (an unobservable meaning) using a self-report questionnaire score (an operational definition). The score would be, at best, an observable manifestation of happiness, but even this cannot be known from an empiricist framework because the relationship between the observable and the unobservable is a non-empirical question (Slife et al., 2005). Likewise, the insistence for operationalization is a non-empirical claim; at best, it must rely on a logical or theoretical argument that can never be confirmed nor denied by empirical evidence alone.

How, then, would one evaluate whether an unobservable meaning is connected to an observable operationalization? Here we can take cues from the common-factors movement’s denunciation of the atomistic assumption—we must appeal to a wider...
framework of evidence, one that includes the study of unobservable meanings. As I will demonstrate, reputable methods—qualitative methods—exist which intentionally avoid operationalization and involve the study of unobservables. Thus, if unobservable meanings are in fact a subject of interest for psychological practice, then the proper approach is to turn to qualitative methods “rather than to turn unobservable meanings into something they are not” (Slife & Wendt, 2005).

The APA policy’s insistence for operationalization highlights just one way that the common-factors framework is driven by the constraints of empiricism. It is discussed as an example here to demonstrate how psychologists are interested in non-empirical content, but are nonetheless restricted to an empiricist epistemology in method (Slife et al., 2005). This restriction causes the reshaping of unobservable phenomena to fit empiricist methodology, without considering whether alternative method philosophies are more appropriate—just as the conception of comorbidity is inappropriately reshaped by an EST framework.

Marginalization of Non-empirical Philosophies and Methods

Just as an EST framework ignores the fact that certain conceptions of evidence are outside the realm of RCTs, so does a common-factors empiricist framework ignore the fact, established earlier, that unobservable phenomena are inherently—now and forever—outside the realm of empirical methods. Like an EST framework, this mistake commits a preinvestigatory bias against alternative conceptions of evidence.

For example, in its report’s discussion of appropriate methods for EBPP, the APA task force appears to ignore the existence and potential of non-empirical methods. This can be seen in the task force’s evaluation of qualitative methods. Although commonly misunderstood as originating from an empiricist methodology, qualitative methods are based on an alternative philosophy of science that neither requires nor prefers operationalization (Slife et al., 2005). Early qualitative researchers were interested in unobservable phenomena themselves, not their supposed manifestations. Therefore, they have developed alternative, qualitative methods that are better suited than empirical methods to understand and investigate these meanings, including their connection to observable experiences (Slife & Wendt, 2005).

Although the task force’s report includes qualitative research on its list of acceptable methods, it fails to understand and value qualitative research as a non-empirical method. When one considers, for example, the report’s insistence for operationalization, it is puzzling to know how exactly non-empirical qualitative methods would inform evidence-based practice. The report is hardly informative on the matter, stating merely that “qualitative research can be used to describe the subjective, lived experiences of people, including participants in psychotherapy” (APA, 2006, p. 274). How do descriptions of “subjective, lived experiences” inform evidence-based practice? From the task force’s report, the answer is unclear—one can only surmise, given the policy’s requirement of operationalization, that qualitative research could lead to the development of new or improved operationalizations that can then be isolated, investigated, and implemented for evidence-based practice. Such a view is driven by an empiricist epistemology, causing a distorted and marginalized conception of the role of qualitative research.

Another clear signal that the task force misunderstands and misrepresents qualitative research—and that also indicates the APA policy’s empiricist framework—is the use of the word subjective in describing the purpose of qualitative research. Of all the methods the task force recommends, the word subjective is reserved only for qualitative research, implying that all other recommended methods are “objective.” In the midst of a discipline that champions objective inquiry, a relegation of being subjective is a second-class citizenship at best (Slife, 2006). A relegation of subjective, in this case, makes sense
only from an empiricist framework; from a non-empiricist framework, empiricist conceptions of "objective" and "subjective" are largely irrelevant (Slife, 2006). This is because an empiricist's definition of "objective" is generally synonymous with "empirical." From a non-empirical perspective, however, the subject matter of qualitative researchers is hardly subjective. The qualitative researcher is not interested, for example, in investigating a subjective interpretation of "love"; she is interested in studying the objective, non-sensory experience of love itself. Therefore, although the task force recommends the use of qualitative methods for EBPP, it does so only in a way that is grounded in empiricism, thereby distorting the nature and history of qualitative research.

That the APA policy misunderstands and misrepresents qualitative research calls into question whether it truly "acknowledge[s] the valid points from all sides of the debate" ("A presidential," 2005, p. 59). Instead, the policy is committed to an empiricist epistemology that causes it to have a preinvestigatory bias against non-empiricist philosophies and the methods they imply. This bias can lead to the misinterpretation and marginalization of a given method—as is the case for qualitative methods—or it can exclude the method altogether, before investigation even begins.

In summary, the common-factors movement and APA policy are driven by the epistemology of empiricism, and this commitment is analogous to an EST framework's commitment to RCTs. Because the policy does not consider whether an empiricist framework is consistent with the nature of real-world patients and circumstances, it reshapes, marginalizes, or ignores non-empirical phenomena and methods in order to adhere to the logic of empiricism.

A Broader Framework: Objective Methodological Pluralism

Thus far, I have argued that neither the EST movement nor the common-factors movement has produced a framework that matches the APA task force's ideals of objectivity and diversity. Such ideals are prevented by the fact that each movement constitutes a preinvestigatory bias against a broader, and necessary, conception of evidence. Thus, the discipline is still in need of a proper framework for evidence-based practice—one that is broad enough to include not only the investigation and implementation of research methods and types of practice, but also their underlying epistemologies. Such is required to avoid the error of uncritically limiting acceptable evidence to a preconceived method or epistemology.

The remainder of this paper will discuss a potentially successful framework, objective methodological pluralism (OMP). As evidenced by its name, OMP is concerned with being "objective," rather than with (subjectively) shaping the world to fit the demands of a chosen method or methodology. It is important to understand, however, that for OMP the meaning of objective is more broadly understood than it is from an empiricist framework. The objectivity of empiricism is often simply equated with an adherence to empiricist methodology, which is assumed—rather than known—to be the objective "reveler of the actual and the truthful, not the bearer of some ideological or economic methodology" (Slife et al., 2005, p. 83). This conception of objectivity is inadequate, as is evident when considering the limitations of empiricism; indeed, empiricism itself is both an ideological and economic methodology (Slife et al.).

4. Psychotherapy researcher Brent Slife (Slife et al., 2005, pp. 93–95; Slife, 2006; Slife & Wender, 2005) has articulated OMP in more depth than will be included in this paper. Here I wish to briefly discuss how OMP promises to be a genuinely objective and diverse framework for evidence-based practice.
In contrast to an empiricist framework, OMP's objectivity is not due to an adherence to a single methodology, but from its aim to be guided by "the truth of our practical experience" (Slife, 2006). In other words, OMP seeks to be driven by the "object of study" (Slife), not a preordained understanding about how to best measure such (Slife et al., 2005). Only by being object-driven, as opposed to method- or methodology-driven, can a framework for evidence-based practice best avoid an ideological or economic methodology.

To be object-driven, OMP requires one to first consider the nature of the object of study, and then utilize or invent the research method or type of practice that is most compatible with the object's nature (Slife et al., 2005). This would require determining, for example, whether the object of study is strictly speaking, a sensory observation (e.g., heart rate, behavioral habits) or a non-sensory meaning (e.g., empathy, therapeutic relationship; Slife, 2006). If "the postulated characteristics of the object" (Slife et al., p. 94) are knowable through sensory observation, then traditional empirical methods might be the preferred route (Slife & Wendt, 2005). However, for the study of unobservable meanings—which, again, constitute a vast portion of the phenomena of interest for evidence-based practice—qualitative methods would be more appropriate.

In examining the object of one's study, researchers and practitioners might ask the following types of questions: What is the nature of the object? Is it a phenomenon that is knowable through empirical observation alone? Or is it an unobservable meaning? If an unobservable meaning, does it also consist of observable manifestations? If so, how are such manifestations related to the unobservable meaning? From an OMP perspective, such questions would be asked before committing to a preconceived method, philosophy, or ideology. Such an approach differs starkly from an EST or common-factors approach, both of which depend on a commitment to a preconceived method or epistemology. This dependency prohibits each framework from considering epistemological, philosophical, or value questions in the first place.

An example will illustrate how the OMP and common-factors approaches might compare—and how OMP best avoids being driven by a preconceived method or epistemology. Suppose that a researcher wishes to examine the role of a therapist's "empathy" in effective psychotherapy (a proven factor of effective therapy, according to the common-factors movement) (Bohart et al., 1998). From the onset, the common-factors approach is constrained to begin with an operational definition of empathy, according to "appropriate scientific methodology" (APA, 2006, p. 274). To establish an empirically demonstrable definition, the researcher would determine certain variables that are assumed—rather than known—to relate to empathy and then determine their efficacy via traditional experimental methods.

An OMP approach, in contrast, would first be concerned with the nature of empathy, rather than automatically adhering to the empirical method and its requirement for operationalization. The researcher might ask, "What am I really interested in?" In this case, she might determine that one cannot assume that empathy is, at bottom, a sensory observation. She might consider, of course, constructing an empirical operationalization of

5. One might prefer the term "subject-driven," "phenomena-driven," "question-driven," "concept-driven," or "patient-driven" (Stephen Yanchar, personal communication, June 26, 2006). There is no problem with these terms, but "object-driven" (Slife's term) will be used throughout this paper to emphasize the objectivity of an object-driven approach.

6. Of course, these types of questions should be considered and reconsidered throughout the process of one's study. An important feature of OMP is its "active and ongoing dialogue about the method values needed to illuminate the objects of inquiry" (Slife et al., 2005, p. 94). In this respect, "various informal investigations and methods could aid in making these decisions, and some value systems even 'try on' to see how helpful they are" (p. 94). The bottom line is that such value systems and methods "would themselves be continuously on trial" (p. 94).
empathy (e.g., vocal properties of therapists that are judged by their patients to have empathy). But she decides against such, determining that she is not interested in manifestations and subjective reports of empathy—she desires an objective study of empathy itself. If her study is to include empirical, operational designs, then it must also include a rationale for how such operations relate to empathy. Such hypothetical relationships, the researcher decides, are inherently unobservable; they lie beyond the realm of empirical methods, because the object of study itself—empathy—is an unobservable meaning, not an empirically demonstrable entity.

With this realization, our hypothetical researcher can do no more if she is trained solely in empirical methods—just as a nails-only carpenter is at a loss when it comes to screws. With expertise in qualitative methods, however, she might be able to determine or create a method that is suitable for examining unobservable meanings. Such a study might include empirical factors, (e.g., practitioners’ vocal properties or patients’ Likert-scale judgments) but they would not be automatically equated with empathy but seen as mere parts or manifestations of a larger, unobservable meaning. Thus, a qualitative approach would need to involve and examine a theoretical connection between such empirical manifestations and the unobservable meaning of “empathy,” as well as examine the relationship between “empathy” and successful practice. This would require, of course, an objective approach to the study of “successful practice”—another unobservable meaning, and one that is “rarely considered in psychology” (Slife, 2006).

Now, it is important to note that OMP does not pretend to be free of assumptions or biases. At bottom, OMP is a pragmatic framework, based on a philosophy that “takes its cues from the practical context of research rather than the abstract propositions of epistemology” (Slife et al., 2005, p. 94). This type of pragmatism is founded, in many respects, to the tradition of William James (p. 94). I do not expect that OMP’s underlying commitment to pragmatism would worry many practitioners in the discipline, especially considering, once again, that the vast majority are at least somewhat pragmatically eclectic or integrationist (see above). Moreover, OMP’s pragmatism should not be terribly troubling to both researchers and practitioners who value the APA policy’s requirement of adapting therapy to the practical needs of clinical circumstances and individual patients. Indeed, the APA policy itself is also underlain in a commitment to pragmatism—how else can one understand its claim that EBPP unlike ESTs, “starts with the patient and asks what research evidence . . . will assist the psychologist in achieving the best outcome” (APA, 2006, p. 273).

The difference, however, between OMP and EBPP is that OMP is ultimately grounded in pragmatism, whereas EBPP is ultimately grounded in empiricism. Therefore, when EBPP claims that it “starts with the patient,” what it really means is that it starts with the patient in a way that presupposes the sufficiency of an empirical methodology. Thus, its conception of the very meaning of “patient” and what it means to “start with the patient” are already grounded in an empiricist framework. Thus, from the very beginning, unobservable meanings of “starting with the patient,” including the methods and practices they might imply, are never considered, in spite of the fact that an attention to

7. It is fair to wonder, at this point, how exactly the study on empathy would be conducted, and I confess that my description is limited in its concreteness. It is beyond the scope of this paper to articulate or create an in-depth example of a particular method for such. Instead, I have attempted to briefly discuss the general principles that might be incorporated in order to maintain an object-driven enquiry. In keeping with the fact that OMP requires an object-driven approach, I cannot even attempt to provide a plausible example without conducting an actual study. This may seem to be an enormous limitation, but actually it is consistent with the overriding claim of this paper: In order to be a truly objective and diverse discipline, psychology cannot determine in advance a detailed method or methodology; instead, its methods must merely be an outgrowth of “the truth of our practical experience.” Because of this, a description of OMP cannot provide details for a prescriptive methodology without falling into the trap of being led by a predetermined method or methodology.
such meanings are a logical extension of “the truth of our practical experience” (Slife, 2006), as argued previously.

Unlike APA's conception of EBPP, OMP is open to the entire range of real-world clinical phenomena, both sensory observations and unobservable meanings. This openness allows OMP to be a truly diverse framework for evidence-based practice—it is not inherently biased against any method or epistemology that is useful for evidence-based practice. Instead, OMP seeks only to be pragmatic, open to whatever methods and epistemologies are a practical outgrowth of “the truth of our practical experience” (Slife, 2006).

Because the truth of the experience of real-world practice is concerned with unobservables, it is therefore necessary for OMP to be concerned with unobservable methods, as well as the epistemologies that guide those methods. This pragmatic openness prevents OMP from being driven by a restrictive epistemological or methodological ideology (Slife et al., 2005). As a consequence, OMP is a more diverse framework than an empiricist common-factors conception of evidence.

Conclusion

When compared to an EST framework, the common-factors movement and the APA policy are a step forward for evidence-based practice toward its ideals of objectivity and diversity. Upon close evaluation, however, it is evident that the APA policy makes similar mistakes. Both are based on limited, biased conceptions of evidence that constitute a preinvestigatory bias against alternative conceptions and the methods they imply. An EST framework is built upon a medical-model of evidence that is biased toward RCT-based treatments for specific disorders, causing it to fail to appreciate and consider alternative research methods and types of practice that are built upon non-medical-model conceptions of evidence. Analogously, the common-factors movement and APA policy are built upon an empiricist model of evidence that is biased toward sensory observable criteria of evidence, causing them to fail to appreciate and consider non-empirical methods and practices, such as qualitative research. Ultimately, objective methodological pluralism (OMP) is a more promising framework for matching the APA task force's criteria of objectivity and diversity.

References


2007 Submission Guidelines

Call for submissions for the Fall 2007 issue of
Intuition: BYU Undergraduate Journal of Psychology

Submissions must adhere to the following guidelines:

• The author (or first author) must be an undergraduate psychology major/minor at a BYU campus during the time he or she wrote the submitted work.
• Articles submitted for publication cannot have been accepted for publication elsewhere.
• Articles (except creative works) must be at least 1,000 words in length and must conform to APA style.
• An electronic copy of all articles must be submitted (see below for further directions). Preferred format for the electronic copy is Microsoft Word. All graphics or photos must be in black and white and of high resolution (300 dpi).

Types of Submissions

• Brief and extended reports of theoretical development or original research (or both). We accept submissions from any field of study in psychology.
• Creative works (psychologically themed poetry, visual art, and personal stories).
• Topical reviews, book reviews, and essays.

Important Information

• Submissions are accepted and processed year-round. The submission deadline for the Fall 2007 volume will be January 15, 2007.
• Those who wish to submit manuscripts, reviews, or creative works to be considered for publication should send an e-mail, with an electronic copy of their work attached, to byupsychjournal@gmail.com and (for BYU Provo students only) place a hard copy of their work in the publications drop box on the third floor of the HBLL (past the general reference desk, in the Publications Lab next to the No Shush Zone).
• Submissions will undergo a peer and faculty review before an offer of acceptance is extended.
• If your submission is accepted, you will be expected to work in collaboration with our editorial board. Publication in the journal is ultimately contingent upon your willingness to contribute to the revising and polishing of your own work.
• Any questions, comments, or concerns should be directed to the Editor-in-Chief at

Editor-in-Chief
Intuition BYU Undergraduate Journal of Psychology
SWKT 1001
Provo, UT 84602
byupsychjournal@gmail.com

Additional information can be found on our Web site at http://intuition.byu.edu/index.html