Editorial: Research, Profoundly Important Reading. and Academic Street Fights

In 2002, when the *Journal of Social Work Values and Ethics* was still in the planning stages, the Editorial Board discussed the desire to offer a special issue on research ethics. I have been teaching research methods since 1980 and statistics since 1992. Thus, the special issue on research is particularly near and dear to me. Between 2002 and 2005, I have been introducing the possibility of this special issue at every appropriate venue in which I had an opportunity to speak (or write in the case of listservs). In one of our listserv transmissions, Judy Krysik had expressed an interest in becoming the special editor for this issue. Judy has been an editorial board member since the fall of 2005, and for the past three years has worked with me as Assistant Chair of Council on Social Work Education's Values and Ethics Track (2007 to 2010). In 2011, she became the Chair of this critically important committee.

Judy's performance as editor has been nothing less than stupendous! I think that many professors will be using this special issue for required reading in research and statistics courses for years to come. As Judy and I are both "empiricists," I envision that we will be monitoring the number of hits each article receives. The advantage of an online journal is that the editorial staff can track the frequency of article usage. We can provide the authors an estimate of the number of times their work has been read! For this issue, I know that my students will be providing many of the hits.

The most profoundly important book I have read within that last five years has to been Trudzik's *If You Tell ... It Will Kill Your Mother*. I strongly encourage child welfare practitioners to read this book. Many clinical insights will be gained. Child welfare professors need to include this fine book as part of their reading assignments. To learn why, read the book review within this issue. Normally, I would not discuss a book review in an editorial – except, I learned that Mrs. Trudzik was Judy Krysik's fourth grade teacher! Talk about a small world.

Lastly, regular readers will be familiar with the continuing debate between Spano/Koenig and Adams. The debate started in third issue of the fourth volume with an article entitled, "What is sacred when personal and professional values collide?" [http://www.socialworker.com/jswve/content/view/69/54/] and has been continuing ever since. This debate is the closest thing to a street fight that one can find in academia. The street fight continues in this issue!

Thank you Judy for such a GREAT issue!!!

Stephen M. Marson, Ph.D.
Senior Editor

Editorial Comment: From the Guest Editor

Five timely and thought provoking papers have been compiled for this special issue of the JSWVE. These papers have the potential to stimulate renewed debate and discussion of research ethics in our classrooms and conferences.

Alan Barsky, for example, goes beyond the traditional rules-based approach to promoting ethical research that is presented in most texts and ethics trainings. Instead, he promotes an approach to the ethical conduct of research that is based on the promotion of virtues. Barsky's paper warrants discussion of: the contrast and merits of the rules-based and virtues ethics approaches, the particular virtues that we should pursue in the ethical conduct of social work research, as well as the means to implement a *virtues* ethics approach.

Ethical issues related to methodological issues are presented in two papers. Patrick Dattalo discusses important considerations in sampling that are relevant to our current era of evidence-based practice and the promotion of randomized controlled trials as the gold standard of research. Dattalo's paper presents alternative research designs and rationales for their use. Martin Bloom, the well-known advocate of single system designs in social work introduces the concept of *evaluation-informed specific practice* and discusses it in the context of evidence-based general practice and client-centered practice.

The final two papers take a critical look at populations that are largely invisible or neglected in social work research. Kelly Jackson outlines considerations for social work research with the growing population of multiracial individuals. Alyson Shapiro and I examine the coverage of fathers in family-focused social work research and call for their inclusion not only as a topic, but as actual participants in research.

Thank you to the editors of JSWVE, Stephen Marson and Jerry Finn, who saw the need to provide a forum to discuss ethical issues in social work research. Thank you also to the contributing authors for their creativity and patience in the review process, and to the many reviewers who I relied upon to provide hours of service. Finally, to you the reader, I hope that you will be part of the ongoing dialogue in these areas. If you have a perspective to share, please write to me or the journal editors.

Judy Krysik, Arizona State University
Guest Editor

A Reply to Adams: The Delicate Balance between Private Viewpoints and Professional Duties

Rick Spano, Ph.D., LSCSW and Terry L. Koenig, Ph.D., LSCSW
University of Kansas
Email: ricks@ku.edu

Abstract

In our second response, we reiterated that all personal values must be mediated through the Code of Ethics (Spano & Koenig, 2009). At this time, we expressed concern about Adams' unwillingness to distinguish between free speech in the public square and professional duties. We appreciate Adams' (2009) most recent response and for his willingness to engage in dialogue with us around a broad range of issues associated with our original article entitled, *"What is sacred when personal and professional values collide?"* (Spano & Koenig, 2007/2008). We remain concerned about Adams' continued lack of response regarding the distinction between free speech in the public square and professional duties. In this third response, we discuss the following: (1) middle ground; (2) knowledge and values in social work; and (3) the personal is always political.

Key terms: Code of Ethics, personal is political, ideology, virtue ethics

A Reply to Adams: The Delicate Balance between Private Viewpoints and Professional Duties

1.0 Introduction

The Code of Ethics (NASW, 1999) is written in such a way that it allows for latitude in the translation of its principles and values into professional behavior. No Code of Ethics can be useful if it is written at a level of specificity that attempts to remove or replace individual professional judgment. It is impossible and unproductive to write a Code that attempts to prescribe specific behaviors in all practice situations due the fact that no two situations are fully identical. However, we are not suggesting that ethical guidelines be ignored for they indeed set parameters in which professional behavior can be evaluated. The following points provide areas of agreement and disagreement with what we understand to be Adam's line of argument.

2.0 Middle ground

We do share an area of agreement or a middle ground with Adams. Each of us has written articles that speak to a more common view of what needs to happen in social work education and practice. The focus on developing social workers' character (virtue) through self-awareness, self-reflection and critical thinking is something we share in common with Adams (Adams, 2009; Spano & Koenig, 2003; Spano & Koenig, 2007/2008). From our perspective, the problem is that many of our students do not come into social work programs and the profession having already developed these skills necessary for using something like "virtue ethics." Their interests are often in developing mastery of techniques and strategies for implementing models. Education, in general, and specifically social work education has moved toward teaching models rather than teaching self-reflective and analytical skills necessary for critiquing and thoughtfully applying models.

We agree with Adams' assessment that although virtuous character is important in social work education, current social work educational structures may not lend themselves to cultivating this character which then can be used in the application of knowledge and values in practice behaviors. As noted above, social work education's current emphasis on the technical skills (Schön, 1983) needed for practice

make it difficult to create opportunities for fostering virtues such as perseverance, fortitude and humility. This is especially true in the current context in which resource starved agencies are barely able to meet CSWE standards for minimal supervision.

3.0 Knowledge and values in social work

Adams (2009, par. 5) states that "what is alarmingly symptomatic of the social work academy is the substitution of ideology for evidence, the refusal even to look at evidence to the great detriment of those we teach and the people with whom they subsequently work." In the social sciences, the distinction between knowledge and values is a false dichotomy. Ideology, which permeates all of the social sciences, is the combination of knowledge and values. To suggest that knowledge exists in a value-free environment, presumes that values do not shape the questions we ask, the evidence we collect, and the meaning we assign to that evidence, which in turn, informs our actions. Part of the problem with asserting that social work practice is knowledge based is that we have inadequate knowledge and what knowledge we do have is frequently conflicting. This creates difficulties in addressing practice situations which are typically complex in nature. The choice of which elements of knowledge meet the standard for "evidence" is often based on preexisting personal values. Every social worker brings personal values and perspectives to the table. What we often do is choose evidence that supports our preexisting values or beliefs. As the writers of this third response to Adams, we readily admit that we also have found evidence that bolsters opposing values and beliefs regarding the protections of marriage. For example, some literature points to evidence that married women are at greater risk for domestic violence (i.e., homicide) than women who are not married. Dugan, Nagin, & Rosenfeld (2003) state that, "because most intimate partner killings involve married couples, perhaps the most crucial factor in reducing intimate partner homicide has been the sharp drop in marriage rates among young adults during the past 25 years" (p. 22).

On the other side of the argument, Adams cites literature which supports the protective factors of traditional marriage between a woman and man (e.g., decreased child poverty; and married women as less likely to experience domestic violence). What constitutes "evidence" is shaped, in part, by the values of the people who collect and interpret the evidence. Social work practice is not solely guided by a perspective that emphasizes knowledge; it is shaped by the values held by those who produce and consume the evidence. In these instances, when there is conflicting evidence or a lack of evidence, how do we determine our ethical obligation to clients? Our answer is that social workers should attempt to evaluate the consequences of their actions through an examination of the duties flowing from the principles in the Code of Ethics rather than acting on their own personal beliefs and values to guide their decision making. If we do not acknowledge that there is a common framework (Code of Ethics) that we agree to as a profession, then, how do we ensure that clients will not be disadvantaged based on the social worker's translation of a particular personal point of view into professional behaviors? Our reading of the Code of Ethics is that it is about the protection of clients, not about the "rights" of social workers to use their personal values in directing their actions in practice.

The following example may be informative to help clarify our position on the role of the Code of Ethics as it relates to social justice. Adams cites one definition of social justice as getting the government "off the back of the armies of compassion" (http://www.centreforsocialjustice.org.uk/default.asp?pageRef=44). We, on the other hand, have emphasized the definition of social justice as promoting the well-being of vulnerable populations (NASW, 1999, p. 5). Who gets left out and who gets served with these competing definitions of social justice? In our professional lives, what should be a guiding principle for choosing from among multiple views of social justice? There is nothing inherently, morally wrong with a definition of social justice based on compassionate conservativism designed to emphasize private charity or Marxist principles designed to emphasize big government. We challenge each group to demonstrate that their view of social justice is consistent with the Code's emphasis on social justice for promoting the well-being of vulnerable populations.

One final illustration highlights some of the problems with only relying on a personal view of social justice in professional practice. If I'm doing community organization in social work, and I want to organize upper middle class white people to help keep African Americans from moving into their community based on their belief that it will harm their property values, am I promoting social justice as a social worker? Where would I go to look for guidance and criteria that would inform professional behavior about whether or not I can participate in such activities as a social worker?

4.0 The personal is always political

In conclusion, if duties are somehow devoid of ideology, how do we arrive what is good? Every

definition of good or virtue has an ideological basis that rests on community values. There are many communities with diverse perspectives and/or ideologies. It is our position that people do not separate ideology (i.e., knowledge and values) from duty or action; these components inform each other. Whether a person is an Evangelical Christian or a radical feminist, his/her ideology informs the understanding of duty, which translates into actions. Our position is that a Code of Ethics represents the current perceptions about the duties and obligations that social work professionals adhere to in their practice. In Adams' latest response, he uses an old feminist slogan, "A woman needs a man like a fish needs a bicycle." Perhaps the paraphrase of another feminist slogan might more appropriate for this discussion, "The personal is [*always*] political," irrespective of the ideology of the person. One's personal ideologies and perspectives *will* bleed through and influence professional behavior. A person can believe whatever he/she wants to believe in private life. We agree with Adams and have never said that a social worker has to pass a litmus test regarding his/her personal beliefs. However, a professional's actions flow from these personal beliefs and values and our Code of Ethics provides an important way of guiding choices that reflect the current agreements about social workers' actions or behaviors.

References

Adams, P. (2009). Code of ethics or ideological club? *Journal of Social Work Values and Ethics*, *6*(1). Retrieved January 21, 2010, from http://www.socialworker.com/jswve/content/view/112/67/.

Adams, P. (2009). Ethics with character: Virtues and the ethical social worker. *Journal of Sociology and Social Welfare, 36*(3), p. 83-105.

Adams, P. (2009). Letter to the editor: Limits of a code of ethics. *Journal of Social Work Values and Ethics, 6*(3). Retrieved February 10, 2010, from http://www.socialworker.com/jswve/content/view/127/69/.

Dugan, L., Nagin, D. S., & Rosenfeld, R. (2003). Do domestic violence services save lives?*National Institute of Justice Journal*, November (250), 20-25.

George, R. P. (2001). *The clash of orthodoxies: Law, religion, and morality in crisis.* Wilmington, DE: ISI Books.

National Association of Social Workers. (1999). NASW *code of ethics*. Washington, DC: Author.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Spano, R., & Koenig, T. L. (2003). Moral dialogue: A worker-client interactional model. *Social Thought, 22,* 91-104.

Spano, R., & Koenig, T. (2007/2008). What is sacred when personal and professional values collide? *Journal of Social Work Values and Ethics*, *4*(3). Retrieved Feb 3, 2009, from http://www.socialworker.com/jswve/content/view/69/54/.

Spano, R., & Koenig, T. (2009). Free speech and special responsibilities of the profession. *Journal of Social Work Values and Ethics, 6(1)*. Retrieved February 10, 2010, from http://www.socialworker.com/jswve/content/view/115/67/.

The Virtuous Social Work Researcher

Allan Edward Barsky, JD, MSW, Ph.D.
Florida Atlantic University

Abstract

Social workers can learn how to *conduct research in an ethical manner* or they can learn how to *be ethical in their roles as researchers*. This article explains the differences between these approaches and articulates what it means to be a virtuous social work researcher.

Key Words: research ethics, virtues, ethics of care, researcher traits

1.0 Introduction

Most social work research textbooks include at least one chapter on research ethics (Grinnell & Unrau, 2005; Rubin & Babbie, 2008). Typically, these educational materials teach students how to apply the relevant legal rules, ethical principles, and standards of practice in research situations. For instance, students learn how to apply the federal guidelines under the Common Rule (2005) for research involving human subjects. They also learn how to apply the principles and standards from the NASW Code of Ethics (2008) concerning informed consent, confidentiality, integrity, and protecting the rights and wellbeing of vulnerable research participants (Standards, 1.03, 1.07; 5.01, and 5.02). Without explicitly stating the ethical theories behind these approaches, these textbooks are incorporating deontology (the study of ethical duties) and teleology (the study of behaviors and their ethical consequences). Unfortunately, many research textbooks and courses ignore or play down the potential of a third approach, virtue ethics (Pring, 2001). Whereas deontology and teleology focus on ways of thinking and making rational choices about ethical behaviors, virtue ethics focuses on the development of good (or virtuous) character states

(Barsky, 2010; Dolgoff, Loewenberg, & Harrington, 2009). Thus, virtuous researchers are investigators who are disposed to incorporating good values, morals, and ideals in all elements of their being (Murphy, 1999). They encapsulate virtue in their professional personae. They are not moral thinkers, but moral agents and beings. They are not simply motivated by social pressures to follow rules or the fear of punishment if they do not. Virtuous individuals act ethically because they are internally motivated to do so (Cohen & Cohen, 1999). This article proposes core virtues that define the virtuous social work researcher (SWR) and demonstrates how virtue ethics can be nurtured in the context of social work research. The virtues proposed in this article are intended to open a dialogue among SWRs regarding the virtues to which they aspire. They are not intended to be conclusive.

The following analysis begins with an overview of virtue ethics, clarifying ways in which this approach differs from a traditional approach to making ethical decisions through the rational application of rules, principles, and standards. The subsequent sections offer suggestions on the core traits of virtuous SWRs, including virtues that are common to all SWRs and others that depend on the type of research (e.g., quantitative, qualitative, or action research). The concluding section offers suggestions for teaching and nurturing virtues among social work researchers. The United States Army has a slogan, "Be all you can be." Perhaps a variant of this slogan could be used as a call to virtue among social worker researchers.

1.1 The Meaning and Import of Virtue

Aristotle, the ancient Greek philosopher, believed that all people share an ultimate human end or purpose. He suggested that a good or virtuous life

can be defined by how well each person fulfills his or her natural purpose (Cohen & Cohen, 1999; Knight, 2007). Eudemonia, living the good life, means incorporating virtue in all aspects of oneself: thinking, behaving, feeling, and being. Virtues are not simply rules or choices, but rather, enduring and transcending character traits (Boatright, 2006). Thus, a person who lives by the virtue of kindness should not only think kind thoughts, but also act kindly, emote kindness, and be kind in every dimension of his or her life.

For Aristotle, one of the overarching virtues is moderation, avoiding excesses (Parrott, 2006). Although one should not be too shy, one should also avoid being too shameless. Thus, one should be proud and self-confident, but not arrogant or brazen. Similarly, one should generally be relaxed rather than irritable, but not overly calm or blasé. A person who is overly relaxed may not be inspired to confront a social injustice. A person who is overly irritated by social injustice may react inappropriately, perhaps with violence or other counterproductive responses. Virtue requires balance rather than being "holier than thou" or taking any quality to the extreme (Murphy, 1999)

Although traditional virtue ethicists such as Aristotle, Plato, Confucius, Mengzi, and the Stoics focused on universal virtues – virtues that pertain to all people – modern ethicists have begun to explore how virtue ethics applies to specific professional groups, including educators, psychotherapists and mental health professionals, business people, and the military (Cohen & Cohen, 1999; Walker & Ivanhoe, 2007). The virtues required for one professional context may differ from the virtues required for another. Thus, a core virtue for a court judge is neutrality and impartiality. A judge should not be relational, in the sense of uniting or partnering with one party or another during a court case. In contrast, relationality is a core virtue for social workers. In order to build trust, develop a therapeutic alliance, and work effectively with clients, social workers need to partner with their clients in a collaborative, caring manner (Cohen & Cohen, 1999). That is not to say that bias is a virtue for social workers, or caring is a vice for judges. Rather, each profession may have a different set of core virtues that fits for its role or context of practice (Walker & Ivanhoe, 2007). *Difference in core virtues* does not mean that each profession's core virtues are opposites, just different. Further, *difference in core virtues* does not suggest that one profession's virtues are better or worse than the others. In accordance with the principle of respect for the dignity and worth of all people (NASW, 2008), social workers should not impose negative judgments on individuals or groups who aspire to different virtues.

One of the primary arguments in favor of using a virtues paradigm to instill ethics in researchers is that the existence of ethical rules, obligations, and standards is not sufficient to ensure ethical practice. The Common Rule and professional ethical standards define what types of conduct are appropriate and inappropriate. Mandatory ethics training requirements for researchers ensure that they are informed of their ethical responsibilities. Strict regulations provide for increased accountability, as institutional review boards oversee research proposals and researchers. Thus, researchers know they may be held to account for any ethical breaches arising out of their research. Sanctions for scientific misconduct may include public or private censures, firing, financial penalties, ineligibility for future grants or contracts, and civil lawsuits to compensate those who were hurt by the misconduct (Gibelman & Gelman, 2005). Historically, many of the laws and standards of practice for human subjects research were developed in response to incidents of horrendous research practice: for instance, the Nuremberg Code developed in 1948 was in response to the sadistic experiments conducted by Nazi scientists on Jews and other vulnerable populations, and the National Research Act of 1974 and Belmont Report of 1978 was initiated in response to the Tuskegee research in which African American participants were intentionally denied information and treatment for syphilis (Drewry, 2004). But was lack of legal and ethical guidelines truly the underlying cause of the Nazi and Tuskegee atrocities? Rather, was there something in the training, culture, or social context of the researchers that allowed them to conduct research in a knowingly harmful manner? Might it be more effective to build an educational system and culture that promotes the virtues of integrity, moral courage, compassion, and empathy among researchers who work with human subjects? If we could promote virtues as internal motivators for researchers, would we need to legislate every aspect of what makes research moral or ethical? "No set of principles (and thus no ethical code) can exhaustively shape the moral deliberation which inevitably researchers are caught in" (Pring, 2001, p. 412).

In spite of the development of national and international research laws and standards, scientific misconduct continues to be a problem. The most common forms of misconduct include fabricating the process and outcomes of the research, and failure to protect human participants by giving them incomplete or inaccurate information about the risks of the research (Gibelman & Gelman, 2005). One can

speculate on the motivations for such misconduct, including pressure from universities or other employers to produce and publish research. However, the issue does not seem to be related to the lack of clear research rules, moral obligations, or standards of practice on these issues. Perhaps it is time to provide more emphasis on nurturing virtues (Pring, 2001), supplementing the current focus on teaching how to apply ethical rules, principles, and standards.

1.2 Determining "Social Work Researcher" Virtues

A key challenge in applying virtue ethics is determining which character traits are of primary importance for social workers to flourish as researchers. To guide this process, it may be useful to consider three dimensions of SWR virtues: 1) What are the universal virtues for social workers[1], and what are their implications for social workers as researchers; 2) What additional virtues are imperative for social work in their roles as researchers; and 3) How might the virtues of SWRs depend on the type of research they are conducting?

2.0 Universal Social Work Values

Ideally, the conceptualization of a set of universal social work virtues should be derived from a dialogue between social workers, including social work ethicists. Although the profession of social work has not engaged in discussions of core virtues per se, the profession has engaged a broad range of social workers to develop consensus statements on the definition of social work (International Federation of Social Workers, 2000), and the mission, values, and guiding principles of social work (National Association of Social Workers, 2008).[2] In addition, various ethicists have begun to discern core virtues for social work (McBeath & Web, 2002), psychotherapy (Cohen & Cohen, 1999), psychiatry (Radden, 2007), medicine (Blustein, 2007; Pellegrino, 2007), and related professions (Walker & Ivanhoe, 2007). The following analysis draws from these sources in order to offer a set of core virtues that are vital to all social workers. These are intended to serve as a starting point for discussion, rather than a final statement on social work virtues.

The aforementioned definition, mission, values, and ethical principles (IFSW, 2000, NASW, 2008) suggest that social work is a profession that helps individuals, families, and groups meet their needs and maximize their potential, while also promoting human rights and social justice at community and societal levels. Although these pronouncements do not specifically mention virtues,[3]

they imply that three of the most vital character traits of social workers are the virtues of *caring*, *generosity of spirit*, and *concern for others*.

Caring social workers are ones who attend to the needs of others. Caring exists in the context of relationships and it motivates people to help (Noddings, 2007). Thus, caring social workers strive to understand the needs and perspectives of those they serve, in order to respond in a client-centered manner (Banks, 2006; Vonk, 2000). Caring social workers are particularly sensitive to the vulnerabilities of populations that are affected by social stresses such as poverty, discrimination, and oppression. They do not simply respond to clients or others in a rational, technical manner (Parrott, 2006). They are aware of their emotional responses toward others and they factor in those emotional responses when making ethical choices (Gilligan, 1982). Social workers have a generosity of spirit in the sense that they prioritize service to others. In some instances, caring social workers subjugate their personal interests in order to advance the needs and interests of those they serve. For SWRs, these virtues have many implications:

- When *caring* SWRs make choices about what research interests to pursue, they do so with regard to the wellbeing of others. They may ask, "What types of research promote social justice, human growth, and social development?" rather than, "What types of research will advance my career or make me look good to others?" As altruists, they may take on research interests that are politically unpopular or pay poorly in order to give voice to those who have been disempowered or oppressed. Caring SWRs do not treat research participants as objects or means to an end, but as human beings worthy of dignity and respect (Pring, 2001). Caring SWRs avoid research that would cause harm, even to a small minority. At the same time, caring SWRs understand that it is important to include vulnerable and disadvantaged populations in their research agendas, so that these populations can benefit from research (Antle & Reghehr, 2003).

- SWRs who embrace *generosity of spirit* adopt a humble approach toward the study of biopsychosocial-spiritual phenomena. Rather than assuming that the researcher possesses all the important knowledge and expertise, generous SWRs work with clients and other stakeholders, viewing them as partners or sources of expertise and guidance. They empower others and promote social inclusion because it is the right way for SWRs to relate with others, not because it serves some other purpose.

- SWRs live the virtue of *concern for others* not simply by completing research ethics protocols and forms for informed consent established by their institutions. Virtuous SWRs ensure that research participants are not harmed by the research, even when it means going beyond what is required by the institution. If the research involves risks, concerned SWRs ensure that participants understand the risks and consent voluntarily. Concerned researchers empathize with the research participants (Murphy, 1999). In order to ensure voluntary participation, the SWRs offer participants more than one option. They understand that a participant who has only one option for services may feel pressured into accepting whatever the SWR is offering. SWRs resist self-interested temptations, such as persuading a client to accept certain risks merely because they feel pressure to solicit a sufficient sample size within a short timeframe.

Virtue ethicists suggest three additional virtues for helping professionals such as psychiatrists, physicians, and attorneys: *trustworthiness*, *fortitude*, and *phronesis* (Cohen & Cohen, 1999; Radden, 2007). These virtues are certainly relevant to social work, as they reflect many of the values and ethical principles in the NASW Code of Ethics.

Trustworthiness refers to being reliable, honest, and responsible. Trustworthy social workers provide help only in situations where they are competent to do so. They assume a relationship of fidelity or special care with clients, ensuring that clients receive the services they need and ensuring clients receive services in a safe environment. Trustworthy social workers respect client rights to self-determination, informed consent, confidentiality, and respect for their dignity and self-worth. They act honestly and with integrity. They avoid actions and relationships that put themselves in positions of conflict of interest with their client's needs. They use astute perception, purposeful self-awareness, and prudent appraisal (McBeath & Webb, 2002) to ensure they do not impose their values or beliefs on clients. They accept responsibility and accountability for their actions and continuously strive for self-improvement.

For SWRs, *trustworthiness* means maintaining the faith of the research participants, funders, government, and public. Trustworthiness is particularly important in terms of protecting research participants from risks inherent in their research. Trustworthy SWRs understand the importance of confidentiality from the participant's perspective. They take appropriate safeguards to protect private information and to help participants feel secure in sharing information. Trustworthy SWRs are disposed toward integrity by providing participants with full information about the nature, risks, and benefits of their research. They answer participants' questions with frankness and honesty. Trustworthiness includes being honest with oneself, not just with others (Paul & Elder, 2006). Accordingly, SWRs are true to themselves about actual risks and benefits of their research. SWRs avoid deceptive practices, knowing how deception constitutes a significant breach of trust (Antle & Regehr, 2003). Although limited deception may be ethically justified for some research purposes, they consider whether and how such deception would be viewed from the research participants' perspectives. Would the form of deception proposed for this case cause research participants or the public to view the researchers as undependable or dishonest (Murphy, 1999)? SWRs possess a heightened awareness of the vulnerability of research participants. Given the fiduciary nature of their trust relationships, SWRs adopt the highest levels of transparency and integrity in their work. When they make promises to research participants, they honor their promises. When working with people of color, people with disabilities, or other vulnerable populations, trustworthy SWRs do not simply parachute in to conduct research and then quickly abscond without concern for the impact of their research on the research participants. Trustworthy SWRs safeguard the interests of their research participants whether or not others are watching, asking, or are in a position to discover possible breaches of trust (Pring, 2001). SWRs are modest about their research findings, resisting the temptation to embellish the results in their publications or reports (Pring, 2001).

Fortitude refers to having the moral courage and strength to do what is right (Paul & Elder, 2006). Fortitudinous social workers advocate for the needs and interests of clients and vulnerable populations even when they face challenges such as lack of resources, powerful adversaries, and strong systems that support the status quo. Fortitude means having the muster to act ethically even when the worker risks negative repercussions from clients, employers, government, or others. Fortitude may be viewed as a precondition for other virtues, as virtue invites people to be ethical even when the situation makes it difficult to be ethical (Blustein, 2007).

In a research context, social workers may face many pressures to act in a less than ethical manner. When evaluating the effectiveness of a social program, administrators or other stakeholders may pressure SWRs to find and report positive outcomes. When SWRs are being paid to study a particular phenomenon, they may feel pressure from the payer to produce certain types of findings (e.g., in

support of the payer's political interests). When a college dean asks a student SWR to conduct research but allow the dean to claim sole authorship and full credit, the SWR may feel pressured into compliance. Fortitude means that SWRs resist such pressures to act unethically, even when their jobs, salaries, or degrees are at stake.

Phronesis refers to practical wisdom (Radden, 2007). Social workers do not simply use information and knowledge from research and textbooks. They make use of existing scientific knowledge and theory, but also develop their own understanding of clients and interventions through strategic reflection, evaluation, and critical analysis. The practical wisdom of social workers develops over time, as workers draw from various experiences in their professional and personal lives.

For SWRs, phronesis has implications for research design and implementation. SWRs do not simply rely on textbook information and research protocols for how to design and implement research. They make use of their experience working with clients and research participants to determine how to act in particular situations. General research protocols may suggest, for instance, that researchers should use written consent forms. If the researcher has experience suggesting that a certain population would prefer oral consents (e.g., due to cultural issues or stress factors), the researcher may encourage the institutional review board to allow oral consents (Gordon, 2003). Once an institutional review board has approved a research application, the SWR does not simply follow the research application as written. The SWR continues to use phronesis to assess and respond to any ethical issues that may arise during implementation. Consider, for instance, a research application that requires the researcher to obtain the consent of guardians in order to interview individuals who have been deemed mentally incompetent (i.e., wards). While implementing the research, the SWR discovers that some wards feel restricted in what they can disclose to the SWR, fearing repercussions from their guardians (Kroch, 2009). Using phronesis (practice experience and wisdom), the SWR goes back to the institutional review board to suggest a revised protocol to address these concerns. SWRs who use phronesis understand that they continue to learn about research design and processes as they conduct research.

3.0 Researcher-Role Virtues

Virtues that are particularly vital to social workers in their role as researchers include inquisitiveness, precision, and discernment.[4] *Inquisitiveness* refers to having a curious, questioning disposition. Inquisitiveness is reflected in Standard 5.02 of the NASW (2008) Code of Ethics in relation to research, and in the CSWE (2008) Education Policy and Accreditation Standards in relation to scientific inquiry. Inquisitive SWRs are interested in seeking problems, asking questions, learning, building knowledge, and identifying the unknown. SWRs search for truth or insight (Pring, 2001). They strive to eliminate or reduce misinformation, error, and ignorance. They are motivated to gain a better understanding of individuals, families, groups, organizations, communities, and other social systems. They do not only seek to prove their preferred hypotheses; they are open to learning the unexpected. They are comfortable with research findings that are uncertain or conflicting with their prior knowledge or beliefs. They view research as not simply a means to answering questions, but as a vehicle to identifying new questions and problems. SWRs are open to receiving criticism of their research (Pring, 2001), as criticism leads to further questioning and inquiry.

The virtue of *precision* means concern for accuracy. Precision-oriented SWRs pay attention to detail, including the rigor of their research methods and the accuracy of their data gathering, analysis, and reporting. They guard against sloppiness and mistakes, even when presented with challenges such as limited time and resources. Precision-oriented researchers understand how vigilant attention to validity, reliability, dependability, and related research constructs foster the virtue of trustworthiness (as defined earlier).

Discernment refers to the quality of keen judgment. Discerning SWRs are not simply technically accurate in their work; they attend to issues requiring choice and make use of sound critical thinking in order to determine how to respond. Discernment is vital in many aspects of research, including determining which methods fit best for a particular research question, how a sample should be recruited, how research results should be interpreted, and how to respond to ethical dilemmas. Consider, for instance, an SWR who discovers that a research participant has been threatened by a man she describes as her pimp. The SWR is very concerned about the participant's safety, but the participant says she does not want to go to the police or take other precautions. Using discernment, the SWR considers relevant ethical principles, including individual autonomy, beneficence, nonmaleficence, and justice (Beauchamp & Childress, 2009). The discerning SWR reflects on his[5] own values and beliefs, as well as those of the participant. He explores various options and invites help from others, including the agency's attorney (Barsky, 2010). He does not merely rely on the attorney's legal advice. He makes

sure he understands the basis for the advice and considers whether he should take other factors into consideration.

4.0 Method-Specific Virtues

Although some virtues are applicable for all methods of research, other virtues may be dependent on the specific method or approach to research. This section contrasts virtues that are relevant to quantitative, qualitative, and action research. Note that although this section highlights some of the differences between these methods, the differences should not be exaggerated. Often, the differences are a matter of degree, rather than polar opposites.

Quantitative researchers are disposed toward *impartiality*. When conducting research, they avoid situations of bias or perceived bias. They exercise judgment in an objective manner. Although they often identify hypotheses in the early stages of their research, they plan and implement research in a manner designed to identify the truth about the phenomena they are studying. They do not simply set out to prove their hypotheses. For quantitative researchers, trustworthiness is based on constructing research in a manner that is consistent with the generally accepted standards of scientific, empirical research. Traditionally, quantitative researchers favor being relatively distanced, detached, and independent, in the sense that they prefer not to be aligned with a particular individual, family, group, organization, or community (Danaher, Danaher, & Moriarty, n.d.). Pure impartiality or independence may not be possible, for instance, because a researcher may be contracted to perform a particular study. Still, the researcher strives to conduct and report research in an unbiased manner. Further, a quantitative SWR balances the virtues of impartiality and caring. As a caring social worker, the SWR attends to the dignity and wellbeing of each research participant. Thus, a SWR may need to compromise obtaining an objective (random) sample in order to ensure that vulnerable clients obtain the services they need. On the other hand, many quantitative researchers have relatively little contact, knowledge, or attachment to their research subjects, and may remain quite independent and impartial.

Qualitative SWRs do not have the same disposition toward impartiality as quantitative SWRs. Rather than seeking to identify objective and universal truths, qualitative SWRs often seek to gain understanding from an emic (within group) perspective. This virtue may be described as being *other-centered*. Qualitative SWRs gather information in a manner that reflects the subjective reality of the individuals and groups they are studying. In some instances, such as participant-observation, qualitative SWRs embed themselves in the situation of the people they are studying. For qualitative researchers, the virtue of trustworthiness is not manifest in being an objective scientist-reporter, but in being able to give accurate voice to the concerns, views, or situations of those they are studying.

Action researchers are *collaborative* rather than independent or impartial. They act as partners with their research participants. They empower their research participants to help design the research, including the research questions, sampling, data gathering, analysis, and reporting (Brydon-Miller, Greenwood, & Maguire, 2004). Decisions about research design and implementation may be affected by the interests and political perspectives of the people that the action researchers are serving (Danaher, Danaher, & Moriarty, n.d.). Action researchers are committed to action. Often, they work on behalf of disenfranchised groups, giving them voice and helping them transform their lives (Antle & Regehr, 2003). Action researchers may specifically design research to influence public policy formation and law reform (Barsky, 2009). Because action researchers are committed to advancing social justice, empowerment, and personal development, impartiality and independence are not their highest values. Action researchers maintain trustworthiness by being open to criticism about the limitations of their research methods and their affiliation with the research participants. Although action researchers are motivated to help others, they act with integrity and honesty. Consider an action researcher who is helping an African American community confront poverty. The community may ask the researcher to identify the negative impacts of poverty to help them advocate for policy change. When explaining research findings to public officials, the researcher is forthright about her research agenda and affiliation, but also describes what steps she took to ensure the trustworthiness of the research (e.g., use of generally accepted procedures for focus groups and data analysis). The virtuous researcher does not conceal who sponsored or helped develop the research (Danaher, Danaher, & Moriarty, n.d.).

Another interesting contrast between types of SWRs is the extent to which they are formulaic or flexible. Quantitative SWRs tend to embrace formulaic qualities, such as regimentation, orderliness, organization, and logical, deductive reasoning (Patton, 2008). Prior to gathering data, they make specific plans about whom to solicit into their research sample, what tests to administer, what questions to ask, and how to analyze the data (e.g., using what types of statistical analysis in order to support or reject previously determined hypotheses).

Quantitative SWRs strive to maintain consistency, asking the same questions or administering the same tests in the same way to each research participant. They avoid deviations to avoid confounding the data. They understand that in order to claim certain types of relationships between phenomena (correlations, causal effects, etc.), they must adhere strictly to their plan of study. Quantitative SWRs do not take formulism to the extreme, understanding that they are working with human beings. Consider, for instance, a SWR whose initial findings suggest the intervention being tested is harmful. The SWR decides not to wait until the findings are conclusive. Rather, she adjusts or terminates the study in order to reduce or eliminate the risk of harm. By changing the study in the midst of data collection, the SWR may taint the sample and data set. However, being formulaic does not mean staying the course when there are vital interests that require change.

Qualitative SWRs embrace flexibility, not just when there are vital countervailing interests, but as a matter of course, throughout their research. Although qualitative SWRs may begin with a well-defined plan of study, they are amenable to changing course at all stages of the research process. From a quantitative perspective, qualitative research may seem messy or chaotic. Yet this messiness and chaos allows qualitative SWRs to remain open to discovering new problems, issues, and opportunities for learning. Whereas quantitative SWRs embrace positivism (looking for patterns or theories that apply universally), qualitative SWRs embrace post-modernism (looking for local truths, within-group differences, and individual idiosyncrasies). Qualitative SWRs employ flexibility as part of their modus operandi. Consider an SWR studying the experiences of victims of torture (Chambon et al., 2001). Initially, she asks questions related to the ability of victims to trust others. Given the responses of the first few participants, the researcher discovers that the key issue for victims is not *trust* per se, but the *ability to form relationships* with others. In subsequent interviews, the SWR asks about the ability to form relationships, rather than asking about trust. Initially, the SWR planned to interview participants from various regions. Through snowball (nonprobability) sampling, most of the participants ended up coming from the Middle East. The SWR decides to change the focus of the research to victims of torture from the Middle East. Because the qualitative SWR is using inductive reasoning, she does not need to stay focused on her original research questions or sample (Grinnell & Unrau, 2005). Qualitative SWRs view social phenomena in terms of complex, dynamic, unpredictable, and unfolding patterns, stories, and meaning (Patton, 2008).

Accordingly, they understand their role as researchers in terms of flexibility and openness to discovery.

Action researchers may use qualitative or quantitative methods, or a combination of both. Accordingly, the degree to which they incorporate flexibility or formulism into their research partially depends on which method(s) they are using in a particular context. Consider an action researcher who is helping an addictions treatment agency evaluate the effectiveness of its services. Originally, the program staff and clients contract the action researcher to help them conduct a study comparing the effectiveness of motivational interviewing versus 12-step facilitation. The researcher requires flexibility in the initial stages of the research, making use of qualitative methods to help the agency establish criteria for success and possible factors contributing to success (Brydon-Miller, Greenwood, & Maguire, 2004). Eventually, they implement a large-scale, quantitative study using experimental design. This component of the study requires use of standardized (formulaic) measures and data analysis. Even when action researchers are using quantitative methods, however, the research process is iterative: The researcher makes use of a continuous feedback loop, providing information about the research process and findings, and inviting responses from the staff and clients (or other constituencies with whom she is working). If the staff and clients wanted to modify the measures of success, the researcher helps them assess the relative merits of making the changes (flexibility) and staying the course (formulism). Given that action researchers work in partnership with their research participants, they tend toward flexibility more than their traditional quantitative counterparts even when they are incorporating quantitative methods. Action research tends to be process-oriented. Thus, action researchers embrace a certain degree of responsiveness, messiness, uncertainty, and chaos in pursuing the interests and carrying out the wishes of their research partners (Brydon-Miller, Greenwood, & Maguire, 2004).

Some researchers might question whether qualities such as flexibility and formulism are moral qualities; perhaps they simply reflect amoral choices about the most effective way to conduct research. Flexibility and formulism do reflect moral choices, however, in the sense that they give priority to different approaches to generating knowledge and learning about the world. To the extent that quantitative researchers embrace formulism, for instance, they are expressing preferences for knowledge based on large numbers so that it can be generalized (Grinnell & Unrau, 2005). In contrast, flexibility among qualitative researchers reflects preferences for knowledge about the personal

realities or unique experiences of individuals, families, or other social units (Patton, 2008). The differences in preferred virtues does not mean that one approach to research is ethically superior to the other, just different.[6] In fact, Aristotle's virtue of moderation suggests that researchers of all persuasions should be neither too flexible nor too formulaic. Hence, a researcher who primarily conducts quantitative research, should not be too formulaic (rigid), and a researcher who primarily conducts qualitative research should not be too flexible (disorganized). Balancing is required for various contexts of research, and all researchers should respect the value of alternate forms of research that embrace different virtues.

5.0 Conclusion: Future Directions for Virtue Ethics in Research

Historically, when researchers have acted in an unethical manner, the response of governments and agencies has been to develop new rules and ethical standards to govern research. Although this legalistic approach to fostering ethical research has its merits, legislation and policy cannot guarantee the highest standards of ethical conduct. Some people may act unethically by finding ways to skirt the rules. Others may act unethically because they know the likelihood of getting caught is low. Still others may adhere to the minimum standards of conduct that are enforceable by law, but neglect the highest aspirations or ethics of researchers and social workers. This article recommends virtue ethics as a means of promoting research ethics, offering a range of virtues that SWRs can carry with them in various aspects of their work.

Social work education has never been based on a simple model of transferring knowledge to students. Social work education includes providing students with opportunities for experience, reflection, and circumspection (McBeath & Webb, 2002). Thus, a virtues approach fits well for social work education. Both classroom and field education experiences provide students with a social context in which to nurture virtues (Pring, 2001). Thus, social work research education should focus on more than teaching the laws, policies, and standards that govern ethical research practice. Research education should foster inner qualities such as caring, generosity of spirit, concern for others, trustworthiness, fortitude, phronesis, inquisitiveness, precision, discernment, impartiality, other-centeredness, collaborativeness, formulism, and flexibility. Although some pairs of virtues may present researchers with conflict, educators can help researchers learn how to use moderation, balance, and critical thinking to deal with such conflicts. Some critics of virtue ethics argue that it is too idealistic and that it ignores the need to teach about ethical duties, legal obligations, and how to assess the consequences of different courses of action (Walker & Ivanhoe, 2007). However, teaching virtue may be viewed as supplementing other ethics education, not replacing it (Pellegrino, 2007).

Creating a culture of research virtues does not begin and end with educational institutions. If we want virtuous researchers, then we need a virtuous research community (Murphy, 1999; Pring, 2001). Research institutions, associations,[7] conferences, journals, sponsors, and support groups can play a vital role in fostering research virtues. When a researcher violates ethical standards such as confidentiality, informed consent, or protecting participants from harm, professional publications and the pubic media are quick to report this information. Rather than focusing only on ethical failures, research organizations should promote the best of research virtues and celebrate researchers who act as models of virtue: for instance, those who have used moral courage to do what is right in the face of risk and opposition, those who have developed and implemented more effective ways of ensuring that the rights of research participants are protected, or those who have mentored novice researchers with the highest ideals of social work research.[8] In order to foster virtues such as fortitude, trustworthiness, and concern for others, social work research organizations could provide forums for discussing challenging issues, as well as experiential opportunities for developing and testing virtues (e.g., simulations and role-plays). Research organizations could also facilitate mentoring and other support systems for people who may reach out for moral encouragement and guidance (Murphy, 1999).

By focusing on virtues rather than rules, principles, and standards, SWRs may engage in different types of dialogues, encounters, and experiences – ones that permit SWRs to integrate and test their ethical responses, developing emotional intelligence and behavioral skills, as well as cognitive understandings and reasoning. Virtues are lived, not just discussed. What does it mean for a social work researcher "to live the good life" as espoused by Aristotle and more recent virtue ethicists? The possibilities are endless.

References

Antle, B. J., & Regehr, C. (2003). Beyond individual rights and freedoms: Meta-ethics in social work research. *Social Work, 48*(1), 135-144.

Association for Practical and Professional Ethics (n.d.). Retrieved October 11, 2009 from http://www.indiana.edu/~appe.

Banks, S. (2006). *Ethics and values in social work* (3rd ed.). New York: Palgrave Macmillan.

Barsky, A. E. (2009). Social work research and the law: How LGBT research can be structured and used to affect judicial decisions. In W. Meezan & J. I. Martin. (Eds.). *Research methods with gay, lesbian, bisexual, and transgender populations* (pp. 372-401). New York: Routledge.

Barsky, A. E. (2010). *Ethics and values in social work: An integrated approach for a comprehensive curriculum.* New York: Oxford University Press.

Beauchamp, T. L., & Childress, J. F. (2009). *Principles of biomedical ethics* (6th ed.). New York: Oxford University Press.

Blustein, J. (2007). Doctoring and self-forgiveness. In R. L. Walker & P. J. Ivanhoe (Eds.). *Working virtue: Virtue ethics and contemporary moral problems* (pp. 87-112)*. New York: Oxford University Press.

Boatright, J. R. (2006). *Ethics and the conduct of business* (5th ed.). Upper Saddle River, NJ: Pearson.

Brydon-Miller, M., Greenwood, D., & Maguire, P. (2004). Why action research? *Action Research, 1*(1), 9-28. Retrieved October 11, 2009 from http://web.uvic.ca/ocbr/assets/pdfs/Why_Action_Research.pdf.

Chambon, A. S., McGrath, S., Shapiro, B. Z., Abai, M., Dremetsikas, T., & Dudziak, S. (2001). From interpersonal links to webs of relations: Creating befriending relationships with victims of torture and of war. *Journal of Social Work Research,* 2(2). Retrieved October 11, 2009 from http://www.ccvt.org/interpersonal.html.

Cohen, E. D., & Cohen, G. S. (1999). *The virtuous therapist: Ethical practice of counseling and psychotherapy.* Belmont, CA: Brooks/Cole.

Common Rule. (2005). United States Code of Federal Regulations, Title 45, §46 (Authority: National Research Act [1974]. 5 U.S.C. 301; 42 U.S.C. 289[a]). Retrieved October 11, 2009 from http://ohsr.od.nih.gov/guidelines/45cfr46.html.

Council on Social Work Education (CSWE). (2008). Educational policy and accreditation standards. Alexandria, VA: Author. Retrieved October 11, 2009, from http://www.cswe.org/CSWE/accreditation/2008+EPAS+Handbook.htm.

Danaher, P. A., Danaher, G., & Moriarty, B. (n.d.). Risks and dilemmas, virtues and vices: Engaging with stakeholders and gatekeepers. *Australian Traveller Education Research.* Faculties of Education and Creative Arts and Informatics and Communication Central Queensland University, Australia. Retrieved October 11, 2009 from http://www.aare.edu.au/03pap/dan03516.pdf.

Dolgoff, R., Loewenberg, F. M., & Harrington, D. (2009). *Ethical issues for social work practice* (8th ed.). Belmont, CA: Brooks/Cole—Cengage.

Drewry, S. (2004). The ethics of human subjects protection in research. *Journal of Baccalaureate Social Work, 10*(1), 105–117.

Gibelman, M., & Gelman, S. R. (2005). Scientific misconduct in social welfare research: Preventive lessons from other fields. *Social Work Education, 24*(3), 275–295.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Harvard University Press.

Gordon, E. (2003). Trials and tribulations of navigating IRBs: Anthropological and biomedical perspectives of 'risk' in conducting human subjects research. *Anthropological Quarterly*, *76*(2), 299-320.

Grinnell, R. M., & Unrau, Y. (2005). *Social work research and evaluation: Foundations of evidence-based practice* (8th ed.). New York: Oxford University Press.

Kroch, U. (2009). The experience of being a dependent adult (ward) – A hermeneutic phenomenological study (Unpublished Ph.D. dissertation), University of Calgary, School of Social Work, Calgary, AB, Canada.

International Federation of Social Workers. (2000). Definition of social work. Retrieved October 11, 2009, from http://www.ifsw.org/en/p38000208.html.

Knight, K. (2007). *Aristotelian philosophy: Ethics and politics from Aristotle to MacIntyre.* Williston, VT: Polity Press.

McBeath, G., & Webb, S. A. (2002). Virtue ethics and social work: Being lucky, realistic, and not doing ones duty. *British Journal of Social Work, 32*, 1015-1036.

Murphy, P. E. (1999). Character and virtue ethics in international marketing: An agenda for

management, researchers, and educators. *Journal of Business Ethics, 18*(1), 107-124.

National Association of Social Workers (NASW). (2008). *Code of ethics.* Washington, DC: Author. Retrieved October 11, 2009, from http://www.socialworkers.org/pubs/code/code.asp.

Noddings, N. (2007). Caring as relation and virtue in teaching. In R. L. Walker, & P. J. Ivanhoe (Eds.). *Working virtue: Virtue ethics and contemporary moral problems* (pp. 40-60). New York: Oxford University Press.

Parrott, L. (2006). *Values and ethics in social work practice*. Exeter, UK: Learning Matters Press.

Patton, M. Q. (2008).*Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage Publications.

Paul, R., & Elder, L. (2006). *The thinker's guide to understanding the foundations of ethical reasoning* (2nd ed.)*.* Dillon Beach, CA: Foundations of Critical Thinking (http://www.criticalthinking.org).

Pellegrino, E. D. (2007). Professing medicine, virtue based ethics, and the retrieval of professionalism. In R. L. Walker, & P. J. Ivanhoe (Eds.). *Working virtue: Virtue ethics and contemporary moral problems* (pp. 113-134)*.* New York: Oxford University Press.

Pring, R. (2001). The virtues and vices of an educational researcher. *Journal of Philosophy of Education*, *35*(3), 407-421.

Radden, J. (2007). Virtue ethics as professional ethics: The case of psychiatry. In R. L. Walker, & P. J. Ivanhoe (Eds.). *Working virtue: Virtue ethics and contemporary moral problems* (pp. 113-134)*.* New York: Oxford University Press.

Rubin, A., & Babbie, E. (2008). *Research methods for social work* (6th ed.) Belmont, CA: Brooks/Cole.

Society for Social Work and Research. (n.d.). About SSWR. Retrieved October 11, 2009, 2009, from http://www.sswr.org/aboutus.php.

Vonk, M. E. (2000). An ethic of care: A framework for ethical decision making with survivors of sexual assault. *Social Thought, 19*(1), 49-62.

Walker, R. L., & Ivanhoe, P. J. (2007). Introduction. In R. L. Walker, & P. J. Ivanhoe (Eds.). *Working virtue: Virtue ethics and contemporary moral problems* (pp. 1-39)*.* New York: Oxford University Press.

Notes:

[1] In this context, universal suggests that the virtues are relevant regardless of whether the worker acts as a researcher, community organizer, policy advocate, administrator, clinician or other role.

[2] This article draws primarily from the NASW Code of Ethics, though one could also draw virtues from the codes of ethics of other countries.

[3] Values reflect enduring beliefs about what is good or desirable. Ethical principles identify types of behavior which are considered moral, right, or appropriate (Dolgoff, Loewenberg, & Harrington, 2009). Because virtues reflect moral qualities that define the entirety of a person's being, virtues embrace values and ethics, but also go beyond beliefs and behavior to include a person's emotions, spirituality, and nature (Barsky, 2010).

[4] Given that all social workers may be involved in research and evaluation as part of their practice, the virtues in this section may apply to all social workers.

[5] For balance and simplicity, this article alternates use of male and female pronouns for different examples, rather than using the more cumbersome "he or she" and "his or her."

[6] The differences between qualitative and quantitative research have been compared to the differences between a classical French garden (as in Versailles) and a classical English garden. French gardens typically include straight, linear patterns and strong structures that apply across a vast area. English gardens are more chaotic, with hills, trees, flowers, or other components flowing more randomly and in less formal patterns. Both styles of gardens may be beautiful and functional, although in different ways.

[7] Note that an association such as the Association for Practical and Professional Ethics (n.d.) specifically promotes ethical practice and virtues in all professions (including research and social work professions). In contrast, the goals of the Society for Social Work Research (n.d.) do not even mention the promotion of ethical practice or virtues.

[8] Jane Addams is often presented as a model of virtue for social workers, particularly from a practice perspective. Consider, who are the models of virtue for social work research?

Ethical Considerations in Social Work Research with Multiracial Individuals

Kelly F. Jackson, MSW, PhD,
Arizona State University

Abstract

Growing diversity in the U.S. has prioritized social work's ethical obligation to develop specialized knowledge and understanding of culture and its function in human behavior and society. One ethnic minority group that is receiving growing attention in the social sciences is multiracial persons, or persons who identify with more than one race or ethnic group. This population represents one of the fastest growing ethnic minority groups in the United States. The growing presence and visibility of multiracial persons in the US demands that social work researchers critically examine and understand the complexity of identity as it applies to people who identify with more than one race. This article will discuss both past and present conceptualizations of multiracial identity, and the methodological challenges specific to investigations with multiracial participants. This article will conclude with recommended strategies for ensuring ethically responsible and culturally sensitive research with multiracial persons.

Key Words: Multiracial, identity, culturally sensitive research, research methods, ethically responsible research

1.0 Introduction

Growing diversity in the U.S. has prioritized social work's ethical obligation to develop specialized knowledge and understanding of culture and its function in human behavior and society (National Association of Social Workers Code of Ethics, 1999, Sec.1.05; NASW, 2001). This ethical requirement governs not only our practice, but recent initiatives to conduct and disseminate research with ethnic minority populations (U.S. Department of Health and Human Services, 2001). This commitment evokes certain challenges for social work researchers since historically ethnic minority groups have been considerably more vulnerable to stigmatization, exploitation and harm in research (Fisher, Hoagwood, Boyce, Duster, Frank,). Ethical planning becomes a crucial component of research with ethnic minorities, and social work researchers must insure that investigations are not only ethically responsible, but flexible to the culture of its participants (Fisher et al., 2002).

One ethnic minority group is receiving growing attention in social science research: multiracial persons (Shih & Sanchez, 2009), or individuals who identify with more than one racial group (Root & Kelley, 2003). According to the U.S. census, approximately 7 million Americans or 2.4% of the total population identify with more than one race (U.S. Census, 2000). This population has steadily increased since the abolishment of anti-miscegenation laws in 1967 (*Loving v. Virginia*), and now represents one of the fastest growing minority groups in the United States (Shih & Sanchez, 2009). Though racial mixing is far from a new phenomenon in the U.S. (Morning, 2003; Rockquemore, Brunsma, & Delgado, 2009), the increased visibility of multiracial persons in the media (i.e., Tiger Woods, President Barack Obama) have inspired a growing number of people to claim membership in more than one racial group (Rockquemore, Brunsma, & Delgado, 2009). Recent legislation allowing multiracial persons to check more than one race on federal race reporting forms (see Office of Management and Budget, 1997) have made the multiracial population a noteworthy demographic group in the United States (Cheng & Lee, 2009; Shih & Sanchez, 2005). In fact some scholars propose that by the year 2050, one in five persons in the U.S.

could claim a multiracial background (Lee & Bean, 2004; Smith & Edmonston, 1997).

Due to the significant population growth and increased visibility of multiracial persons in the U.S., social workers are more than likely to see a rise in the number of clients and family systems who identify as multiracial (Fong, Spickard, & Ewalt, 1995; Hall, 2001). This heightens the need for social workers to understand how growing up in a complex society that continues to construct race based on historic ideals of distinct racial groups, may be challenging for multiracial persons. For instance, there is limited yet mounting evidence that multiracial youth, in particular are at greater risk than their monoracial peers to use substances, engage in violent behaviors, and struggle with self-esteem (see: Bolland, Bryant, Lian, McCallum, Vazsonyi, & Barth, 2007; Choi, Harachi, Gillmore, & Catalano, 2006; Jackson & LeCroy, 2009; Udry, Li, & Hendrickson-Smith, 2003). Researchers posit that this may be due to stressors associated with navigating a multifaceted identity in a mono-racial focused society (Choi et al., 2006; Samuels, 2009). Unfortunately existing research is unable to capture the more dynamic and fluid processes influencing multiracial identity development, those which may or may not be linked to an individual's risk for developing mental or behavioral health problems (Choi et al., 2006; Shih & Sanchez, 2005).

The growing presence and visibility of multiracial persons in the US demands that social work researchers critically examine and understand the complexity of identity as it applies to people who identify with more than one race (Crawford & Alaggia, 2008). Such knowledge is critical to our field's development of culturally sensitive practice models with multiracial individuals and interracial families (Beneditto & Olisky, 2001; Gibbs, 1998; Morrison & Bordere, 2001; Nishimura, 2004; Wardle, 1991). Unfortunately social work research efforts to understand the identity of this diverse group have been minimal. This is discerning since the profession is known, not only for its ethical obligation to understand culture and its function in human behavior and society (NASW, 1999), but unlike other disciplines, social work has the potential to offer a unique, more inclusive understanding of multiracial identity by utilizing the professions ecological and strength-based perspectives (Jackson, 2009). The dearth of multiracial research may be related to the political and definitional challenges associated with multiracial identification, which include the practice of compartmentalizing persons into separate monolithic racial groups (Edwards & Pedrotti, 2008). Due to this, multiracial research, like ethnic minority research in general, is confounded by certain methodological issues (Root, 1992).

This article will discuss both past and present conceptualizations of multiracial identity, and the methodological challenges specific to investigations with multiracial participants. This article will conclude with recommended strategies for ensuring ethically responsible and culturally sensitive research with multiracial persons. The author recognizes that constructs of race and ethnicity are not static and are often conceptually confusing in social science research (Cokly, 2007), therefore, for the purposes of this article, the author defines the term *multiracial* as individuals who identify with 2 or more different racial heritages (i.e., Black, White, Asian/Pacific Islander, and American Indian/Alaskan Native), which is inclusive of derivative terms such as: *biracial* (individuals who identify with 2 different racial groups), *mixed-race* and *mixed-heritage* (Root & Kelley, 2003).

2.0   U.S. Conceptualization of Multiracial Identity: Past and Present

The social identity of multiracial people is strongly influenced by the sociopolitical context of the U.S., including the stagnant societal belief in the biological existence of separate races (Kahn & Denmon, 1997; Root, 1992b). Certain national events have played a significant role in shaping social conceptualizations of multiracial identity and social science research with multiracial persons. These include: slavery, the legalization of interracial marriage, Census 2000, and the recent election of mixed-race President Barack Obama.

Slavery. Slavery in the U.S. (1654 – 1865), particularly the raping of African female slaves by white males, led to the emergence of multiracial (black/white) individuals. In order to preserve the sanctity of the institution of slavery and protect White masters from having to provide patronage to their half-black offspring, legislation was developed in the 1600's to classify multiracial individuals with African American ancestry as black. This became known as the principal of hypo-descent or the "one-drop rule" (Brown, 2001; Graves, 2004). Around the same time, biological and sociological arguments began to arise portraying multiracial individuals as maladjusted and dangerous degenerates due to their dual polarized heritage (Brown, 2001; Wilson, 1987). This perspective was mostly shaped by false biological claims of the existence of a racial hierarch, placing whites above groups of color, both on a genetic and societal level. For example, social scientists during the late 1800's and early 1900's considered mixed black/white persons less intelligent

and capable than white due to possessing black blood (*Hybrid Degeneracy Theory)* (Brown, 2001). This transformed society's view of multiracial people and lead to the construction of the first multiracial identity theories which classified multiracial persons as "marginal" (e.g., Embree, 1931; Park, 1928; Sommers, 1964; Stonequist, 1937). Multiracial research at this time was often based on speculation and non-representative case histories that strongly emphasized the biased notion that interracial persons suffered from social and emotional problems (Johnson & Nagoshi, 1986). This prompted the rise of stereotypes popularizing the character notion of the genetically, mentally, and morally inferior "*mulatto*."

Legalization of interracial marriage. Following the legalization of interracial marriage in 1967 (*Loving v. Virginia*), the U.S. saw a rise in interracial partnerships and a dramatic increase in multiracial children, also known as the "biracial baby boom" (Root, 1992). This prompted social scientists to again revisit the identity development of multiracial persons. An uprising of theories were proposed in the late 80's and early 90's, attempting to shed light on the racial identity options available to multiracial mainly Black/White children (e.g., Gibbs, 1987, *Model of Biracial Identity Conflicts*; Poston, 1990, *The Biracial Identity Development Model*; Jacobs, 1992, *The Identity Development Model of Biracial Children*). Many of these emerging theories also followed the Eurocentric stage model frameworks of early identity development theorists (i.e., Erikson, 1963, and Cross, 1987). These theories posited that multiracial identity development followed a similar linear path as other minority groups. Using this static approach many of these models pre-assigned and limited social identity options to a choice of either black or multiracial (Rockquemore & Brunsma, 2002). In addition, most of these models were deficit-based, ignoring the contributing ecological factors impacting identity development, and consequently pathologizing the multiracial experience (Poston, 1990). Multiracial identity models developed during this time were either conceptual or based on research with small samples of biracial Black/White children. Despite relying on more pathological models of identity, research produced during this time introduced new conceptualizations of multiracial identity, namely that multiracial persons did not struggle psychologically (Gibbs & Hines, 1992).

Census 2000. The changing sociopolitical climate of the 1990's and the growing percentage of persons who self-identify as multiracial prompted the formation of activist groups advocating the legal right of multiracial persons to claim the racial heritages of both parents. Their efforts elicited the ratification of the U.S. Federal Office of Management and Budget's (OMB) categorizes of race, allowing multiracial persons to check more than one race on federal race reporting forms (OMB, 1997). The 2000 Census marked the official end of the anonymity of multiracial persons in the U.S. (Brown, 2001), making them visible in a country who for the most part had not acknowledged their existence (Chiong, 1998). Prior to and following the 2000 Census, there was a notable increase in multiracial identity research, lead often by multiracial researchers themselves (i.e., Clinical Psychologist Maria P.P. Root). One major priority that emerged from this research was the development of a complex theory of multiracial identity – one that was non-linear and reflective of the numerous contextual variables that may influence identity development (e.g., Hall, 2005; Renn, 2003; Rockquemore & Brunsma, 2002; Root, 1999b; Wijeyeshinge, 2001). Using a more ecological lens, multiracial identity research has introduced new evidence that identities can change across contexts, differ within a family system, and shift over the life course (e.g., Jackson, 2009; Renn, 2003; Rockquemore & Brunsma, 2002; Root, 1998). This has contributed to a more in-depth understanding of multiracial identity, which can include: experiencing prejudice, adopting multiple labels to describe one's racial identity, refusing to disown any part of their heritage, and feeling comfortable in multiple ethnic communities (see Brackett et al., 2006; Buckley & Carter, 2004; Guevarra, 2007; Milville et al., 2005).

Age of Obama. Recently the U.S. has experienced yet another reputable and potentially influential shift in this country's conceptualization of multiracial identity. President Obama, who is multiracial (though he labels himself and is portrayed in U.S. media as African American) has described his mixed-race background as the most important and powerful factor that has prepared him for his role as President of the United States (Radutzky & Devine, 2008). His presence has re-ignited nationwide debates on issues of race and multiracial identity (Hendricks, 2008; Samuels, 2006), including previous notions that multiracial individuals have the right to choose how they identify, and the right to claim membership in both multiracial and mono-racial groups (*Bill of Rights for People of Mixed Heritage*, Root, 2001). This racial paradigm shift challenges social science researchers to contribute new advancements to conceptualizing multiracial identity (Shih & Sanchez, 2009). New research is beginning to emerge that utilizes more sophisticated methods (i.e., secondary analysis of national databases, mixed qualitative and quantitative designs,

multiracial identity instrumentation development, etc.) to generate new insight into the complexities of multiracial identity (see Shih & Sanchez, 2009).

Throughout history the social identity of multiracial persons has challenged legislation on patronage, marriage, and racial classification. There is a growing call for social scientists to examine, more constructively, how multiracial individuals experience and define their identity in a changing society, which up until recently was reluctant to acknowledge their existence. Therefore, ethical planning becomes a crucial factor in future research with multiracial persons, and social work researchers must become familiar with some of the noted challenges complicating multiracial investigations. These challenges will be described in the next section.

3.0 Methodological Challenges in Multiracial Research

The demand to include multiracial participants in research investigations evokes certain methodological challenges since long standing notions of racial categorization (i.e., the One Drop Rule) are still prolific in our social institutions. For instance despite recent changes in federal race reporting standards that allow multiracial persons to choose more than one race (OMB, 1997), many schools and social service agencies continue to force multiracial participants to choose one racial category (Townsend, Marcus, & Bergsieker, 2009). This section will discuss the challenges that arise in research with multiracial participants. These challenges fall into two major methodological categories: sampling and data collection.

Sampling. Obtaining multiracial participants for research studies can be a difficult and complex venture due to the non-random distribution of multiracial people in the US (Root, 2003). According to the U.S. census, the majority of people who identify as multiracial (40%) reside in Western states such as Hawaii (21.4% of total state population), Alaska (5.4%), California (4.7%), and Washington (3.6%) (Jones & Smith, 2001). This makes recruitment efforts in other areas of the country more complicated (Root, 1999). For example, the majority of research involving multiracial subjects has taken place in Western states, such as California, and/or large cities, such as New York City, where there are greater percentages of multiracial persons, and others from diverse ethnic and racial groups (Root, 1992).

Researchers have also relied heavily on snowball sampling to recruit multiracial participants. This can substantially hinder the diversity of experiences and make samples more homogeneous

(i.e., racial mixture, SES, education, age) (Root, 1992). Similarly past multiracial identity research recruited convenience samples of college-age students, which represent a homogeneous age and developmental group (Root, 1999). Finally, advertisement can pose a substantial problem in multiracial research since some multiracial persons do not identify as multiracial and are unlikely to respond to ads that request persons who identify as mixed-race (Root, 1999). Leaving out such individuals may constrain efforts to draw a more holistic picture of multiracial identity. Research describing the multiracial experience based on these constrained samples may be limited in their applicability to diverse multiracial persons who are more intermittently dispersed in homogeneous communities throughout the U.S., who do not necessarily identify as multiracial, and who are not college-age.

Data collection. Qualitative methods of inquiry are employed more frequently in multiracial research (Shih & Sanchez, 2005). This may be related to the methods ability to extract the complexities and intimate details of multiracial identity, in a way that conventional quantitative methods cannot (Strauss & Corbin, 1998). Despite this, means of collecting qualitative data with multiracial samples can introduce bias in research. Specifically, a researcher who relies on past conceptual frameworks of multiracial identity (i.e., psychoanalytic stage-models), or one that is unaware of their own beliefs and values about race, interracial relationships and/or multiracial identity, can pose biased questions in interviews with multiracial persons (Root, 1999). For example, asking a multiracial person questions that explicitly focus on any adjustment or psychological problems they have experienced due to being mixed-race.

A second, less explored area of bias in qualitative studies is the influence of the ethnicity of the researcher and the potential bias this may introduce in interviews with multiracial persons (Root, 1999). Specifically, the ethnicity of the interviewer could potentially influence a multiracial participant's response to interview questions about their ethnic identity (Root, 1999). For example, some participants may feel restraint in sharing their true feelings about a certain ethnic group of people, or about their experiences with a certain community of color, because of the interviewer's ethnicity (Brown, 2001; Root, 1992).

Traditional means of categorizing race (i.e., check one race) continues to pose a problem in quantitative multiracial research. In quantitative investigations this practice occurs both directly (i.e., a survey that asks a multiracial participant to select a

race that they *mostly* identify with) and indirectly (i.e., when race data is redistributed for analysis purposes). Both practices discriminate against a multiracial person's right to self-identify with more than one racial group. Recently there is evidence disputing this common practice in identity development research with multiracial persons. Specifically Townsend et al. (2009) found that forcing a multiracial participant to choose a race on measures introduces discrepancies between the outward identity multiracial persons report, and their desired or chosen identity. In addition, the same authors found that putting pressure on multiracial participants to choose a race subsequently caused decreases in self-esteem and motivation (Townsend et al., 2009). This directly contradicts previous identity assumptions that multiracial persons who identify with only one racial group (namely the minority) would have a stronger sense of self (see Davis, 1996).

Standard instrumentation poses an additional challenge in multiracial research. Researchers continue to utilize measures of racial and ethnic identity that have been criticized as inappropriate for understanding identity development among persons with mixed-heritage (Coleman, Norton, Miranda, & McCubbin, 2003). Such measures are often based on monoracial samples and assume the necessity of a single choice (Root, 1992). For instance, the Multi-Ethnic Identity Measure (MEIM) (Phinney 1992), though a highly reliable and useful instrument to measure ethnic identity, on the surface does not appear sensitive enough to gage a multiracial persons multiple ethnic and cultural associations. For example, specific items on the MEIM require the multiracial participant to reference one ethnic group (i.e., item #1: "I have spent time trying to find out more about my own ethnic group"), which may pose a problem for those persons who simultaneously identify with more than one ethnic group. Employing such instruments in investigations can draw biased conclusions about multiracial identity.

The demand to include ethnic minority participants in research evokes certain challenges since they are considerably more vulnerable to stigmatization, exploitation and harm in research (Fisher et al., 2002; Kazdin, 2003). Multiracial persons are no exception, and in order to expand our understanding of multiracial identity, social work researchers should become familiar with strategies to minimize these challenges in future multiracial investigations. These strategies will be discussed in the next section.

4.0 Recommended Strategies for Multiracial Research

This section will present recommended strategies to design more ethically responsible and culturally sensitive research with multiracial participants. These strategies include: (1) using recent theory to conceptually guide research methodology; (2) sampling more inclusively; (3) incorporating more culturally sensitive measures and instrumentation; and (4) including multiracial persons on research teams.

Strategy 1. Rely on more inclusive, ecological-based theories of multiracial identity to guide study methodology. Future studies should continue the pursuit to understand the complex individual, interpersonal, and contextual factors that interact to shape a multiracial individuals' identity. A number of ecological-based models have yet to be fully explored in multiracial literature (Shih & Sanchez, 2009). These models include, but are not limited to: Hall's (2005) *Biracial Identity Development Across the Life Span Model*; Rockquemore & Brunsma's (2002) *Multidimensional Model of Biracial Identity*; Root's (1999b) *Ecological Framework for Understanding Identity Development*; and Wijeyesinghe's (2001) *Factor Model of Multiracial Identity*. Such models could be used as a guide to understand the interconnected effect individual (i.e., phenotype, self-esteem), interpersonal (relationships with peers and family), and environmental factors (homogeneity of community and school) have on a multiracial person's identity experiences. For instance Renn (2003) applied Bronfenbrenner's (1979) *Ecology Model of Identity Development* to her examination of the identity of multiracial college students. By applying such a model, Renn was able to assess environmental factors influencing the racial identity of her multiracial participants, and offer suggestions of what areas institutions could do to enhance opportunities for supporting multiracial student development (Renn 2003).

Social work researchers are advantaged in the application of ecological-based theories due to our professions reliance on the ecological system's model to understand client problems and ascertain solutions (Hepworth et al., 2010). These models can guide the methodology used in multiracial investigations. For instance, the idea that multiracial identity is fluid and changes over time is justification for longitudinal designs, which could take the form of narratives collected over time, or analyses of a national longitudinal data base (i.e., Add Health). Also newer methodologies to emerge in social work, including the extended case method (see Samuels, 2009) and participatory action research (see Gazel, 2007), may serve useful for extricating some of the

more complex contextual factors including the racial composition and attitudes of the community, which may be impacting a multiracial persons development.

Strategy 2. *Sample more inclusively.* Social work researchers should include more heterogeneous samples of multiracial persons, including persons from diverse ethnic backgrounds and persons from multiple generations (Root, 1999). Specifically the voices of multiracial persons who identify with two or more ethnic minority groups (with no White ancestry), and older adults are needed in multiracial literature. In order to offset challenges related to sampling multiracial populations, social work researchers can utilize recruitment strategies such as leverage sampling: when a multiracial participant recruits their sibling, who often identifies differently, in the study (see Root, 1998); or recruitment through multiracial friendly websites on the internet (i.e., multiracial groups on *Facebook*, multiracial websites such as the Mixed Heritage Center: www.mixedheritagecenter.org). Finally, in order to keep the sampling frame broad and avoid excluding those who do not identify as multiracial, social work researcher can leave identity verbiage out of recruitment material and instead ask for persons "from more than one racial group" (Jackson, 2007). Researchers should weigh the potential costs and benefits of utilizing the aforementioned sampling strategies prior to implementation. For instance, though recruiting siblings and persons from multiracial websites could broaden the sampling frame, it may also limit the generalizability of study findings since such persons may be more demographically similar then different (i.e., age, SES, racial composition, etc.).

Strategy 3. *Incorporate more culturally sensitive measures and instrumentation*. Social work investigations of multiracial identity should allow participants to self-identify as multiracial or, at the very least, check more than one race on quantitative measures. For instance a study sponsored by the National Center for Health Statistics found that multiracial respondents prefer a question format that allows them to self-identify as "multiracial" (Johnson et al., 1997). In addition, researchers interested in using standard measures of ethnic identity should make concerted effort to include measures that are sensitive to persons with multiple heritages, and inclusive of the numerous ways a multiracial person may express their identity (i.e., identify as multiracial, other, with more than one race, or monoracially). New measures have emerged that have been designed for, tested and proven reliable with multiracial samples (i.e., *Multiracial Identity Integration measure (MII),* Cheng & Lee, 2009; *Multiracial-Heritage Awareness & Personal*

Affiliation Scale, Choi-Misailidis, 2003). Such measures may prove useful in future multiracial identity research.

Strategy 4. *Include multiracial persons on research teams.* In order to make sure our research efforts are culturally sensitive to the ethnic minority groups being studied it is important to involve members of that group in the research process (Fisher et al., 2002; Gil & Bob, 1999). In multiracial research, involvement of multiracial persons in both instrument development and data collection is crucial. First it is important to have several multiracial persons and/or experts on multiracial identity help develop and/or review constructed quantitative instruments or semi-structured qualitative interview guides, which have the potential to introduce bias in research designs. For example a recent study by Cheng & Lee (2009) assessing multiracial identity integration, used three experts in multiracial research to proofread and edit their newly constructed instrument. Other researchers have used diverse focus groups, inclusive of multiracial persons, to construct identity measures (see Buckley & Carter, 2004). By allowing such careful review, social work researchers are less likely to develop insensitive surveys or questionnaires based on their own biases or past social scientific assumptions about multiracial individuals (i.e., marginal, psychologically burdened, confused, etc.).

Another culturally sensitive strategy is to use multiracial persons as interviewers to increase the comfort level of multiracial participants and allow them to share more personal and relevant information during the interview process (Root, 1992). This strategy is frequently used in social work research to minimize miscommunication and power imbalances between interviewers and ethnic minority participants (Singh & Johnson, 1998). Social work researchers are beginning to incorporate this strategy in qualitative investigations of multiracial identity by pairing it with efforts to enhance rigor and minimize potential biases associated with multiracial investigators collecting, analyzing, and/or interpreting data on the multiracial experience. These rigor enhancing strategies can include: reviewer triangulation, member checking, and using multiple data coders (see Jackson, 2009; Samuels, 2009).

5.0 Conclusion

This article presented a backdrop and foreground for social work researchers to develop ethically responsible and culturally sensitive research with multiracial persons. Specifically this paper discussed both past and present conceptualizations of multiracial identity, including how national events

inspired changes both in societal perceptions of multiracial identity and social science pursuits to understand it. This paper also presented methodological challenges specific to multiracial research, and concluded with culturally sensitive strategies recommended for future multiracial investigations.

References

Beneditto, A. E., & Olisky, T. (2001). Biracial youth: The role of the school counselor in racial identity development. *Professional School Counseling, 5,* 66-69.

Bolland, J. M., Bryant, C. M., Lian B. E., McCallum, D. M., Vazsonyi, A. T. & Barth, J. M. (2007). Development and risk behavior among African American, Caucasian, and mixed-race adolescents living in high poverty inner-city neighborhoods. *American Journal of Community Psychology, 40*, 230–249.

Brackett, K. P., Marcus, A., McKenzie N. J., Mullins, L. C., Tang, Z., Allen, A. M. (2006). The effects of multiracial identification on students' perceptions of racism. *The Social Science Journal, 43*, 437-444.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Brown, U. M. (2001). *The interracial experience: Growing up black/white racially mixed in the United States*. Westport, CT: Praeger.

Buckley, T. R., & Carter, R. T. (2004). Biracial (Black/White) women: A qualitative study of racial attitudes and beliefs and their implications for therapy. *Women & Therapy, 27,* 45-64.

Cheng, C., & Lee, F. (2009). Multiracial identity integration: Perceptions of conflict and distance among multiracial individuals. *Journal of Social Issues, 65*, 51-68.

Chiong, J. A. (1998). *Racial categorization of multiracial children in schools*. Westport, CT: Bergin & Garvey.

Choi, Y., Harachi, T. W., Gillmore, M. R. & Catalano, R. F. (2006). Are multiracial adolescents at greater risk? Comparisons of rates, patterns, and correlates of substance use and violence between monoracial and multiracial adolescents. *American Journal of Orthopsychiatry, 76*, 86–97.

Choi-Misailidis, S. (2003). Multiracial-heritage awareness and personal affiliation: Development and validation of a new measure to assess identity in people of mixed race descent (January 1, 2003). *ETD Collection for Fordham University.* Paper AAI3098135. http://fordham.bepress.com/dissertations/AAI3098135.

Cokly, K. (2007). Critical issues in the measurement of ethnic and racial identity: A referendum on the state of the field. *Journal of Counseling Psychology, 54*, 224-234.

Coleman, H. L. K., Norton, R., Miranda, G., & McCubbin, L. (2003). Toward an ecological theory of cultural identity development. In D. Pope-Davis, H.L.K. Coleman, W. M. Liu, & R. L. Toporek. (Eds.), *Handbook of Multicultural Competencies in Counseling & Psychology* (pp.38-58). Thousand Oaks, CA: Sage.

Crawford, S. E., & Alaggia, R. (2008). The best of both worlds?: Family influences on mixed race youth identity development. *Qualitative Social Work, 7*, 81-98.

Cross, W. E. (1987). A two-factor theory of black identity: Implications for the study of identity development in minority children. In M. J. Rotheram & J. S. Phinney (Eds.), *Children's Ethnic Socialization* (pp. 117-133). Newbury Park, CA: SAGE.

Davis, L. E. (1996). The new ingredient in the identity of black biracial children. Retrieved on December 20, 2004 from http://news-info.wustl.edu/opeds/opeds96/davisDec96.html.

Edwards, L. M., & Pedrotti, J. T. (2008). A content and methodological review of articles concerning multiracial issues in six major counseling journals. *Journal of Counseling Psychology, 55*, 411-418.

Embree, E. R. (1931). *Brown America*. New York: Viking.

Erikson, E. H. (1963). *Youth: Change and challenge*. New York: Basic Books.

Fisher, C. B., Hoagwood, K., Boyce, C., Duster, T., Frank, D. A., Grisso, T., Levine, R. J. et al. (2002). Research ethics for mental health science involving ethnic minority children and youths. *American Psychologist, 57*, 1024-1040.

Fong, R., Spickard, P. R., & Ewalt, P. L. (1995). A multiracial reality: Issues or social work. *Social Work, 40*(6), 725-728.

Gazel, J. (2007). Walking the talk: Multiracial discourses, realities, and pedagogy. *American Behavioral Scientist, 51*, 532-550.

Gibbs, J. T. (1987). Identity and marginality: Issues in the treatment of biracial adolescents.

American Journal of Orthopsychiatry, 57, 265-278.

Gibbs, J. T. (1998). Biracial adolescents. In J. T. Gibbs & M. Haung (Eds.), *Children of Color: Psychological Interventions with Culturally Diverse Youth* (pp. 305-332). San Francisco: Jossey-Bass.

Gibbs, J. T., & Hines, A. M. (1992). Negotiating ethnic identity: Issues for black-white biracial adolescents. In M. P. P. Root (Ed.), *Racially Mixed People in America* (pp. 223-238). Newbury Park, CA: SAGE.

Gil, E. F., & Bob, S. (1999). Culturally competent research: An ethical perspective. *Clinical Psychology Review, 19*, 45-55.

Graves, Jr. J. L. (2004). *The race myth: Why we pretend race exists in America*. New York: Dutton.

Guevarra, R. (2007). Mexipino: A history of multiethnic identity and the formation of the Mexican and Filipino communities of San Diego, 1900-1965. Unpublished doctoral dissertation, University of California, Santa Barbara.

Hall, R. E. (2001). Biracial sensitive practice: Expanding social services to an invisible population. *Journal of Sociology and Social Welfare, 28*, 23-36.

Hall, R. E. (2005). Eurocentrism in social work education: From race to identity across the lifespan as biracial alternative. *Journal of Social Work, 5,* 101-114.

Hendricks, T. (July 21, 2008). Obama raises profile of mixed-race Americans. *The San Francisco Chronicle*. Retrieved from http://www.sfgate.com/cgi-bin/article.cgi?f=/c/a/2008/07/21/MNGC11PND8.DTL&hw=mixed+race+americans&sn=001&sc=1000.

Hepworth, D. H., Rooney, R. H., Rooney, G. D., Strom-Gottfried, K., & Larsen, J. A. (2010). *Direct Social Work Practice: Theory and Skills (8th ed.).* New York: Brooks/Cole.

Jackson, K. F. (2007). Beyond race: Examining the cultural identity of multiracial individuals. Retrieved October 5, 2007 from ProQuest Digital Dissertations database.

Jackson. K. F. (2009). Beyond Race: Examining the facets of multiracial identity through a life-span developmental lens. *Journal of Ethnic and Cultural Diversity in Social Work, 18*(4), 309-326.

Jackson, K. F., & LeCroy, C. (2009). The influence of race on substance use and negative activity involvement among monoracial and multiracial adolescents of the Southwest. *Journal of Drug Education, 39*(2), 195-210.

Jacobs, J. H. (1992). Identity development in biracial children. In M. P. P. Root (Ed.), *Racially Mixed People in America* (pp. 190-206). Newbury Park, CA: SAGE.

Johnson, T. P., Jobe, J. B., O'Rouke, D., Sudman, S., Warnecke, R., Chavez, N., et al. (1997). Dimensions of self identification among multiracial and multiethnic respondents in survey interviews. *Evaluation Review, 21*(6), 671-687.

Johnson, R. C., & Nagoshi, C. T. (1986). The adjustment of offspring of within-group and interracial/intercultural marriages: A comparison of personality factor scores. *Journal of Marriage and the Family, 48,* 279-284.

Jones, N. A., & Smith, A. S. (2001). *Census 2000 Brief: Two or more races, U.S. Census Bureau.* Retrieved from http://www.census.gov/prod/2001pubs/c2kbr01-6.pdf.

Kahn, J.S., & Denmon, J. (1997). An examination of social science literature pertaining to multiracial identity: A historical perspective. *Journal of Multicultural Social Work, 6*, 117-138.

Kazdin, A.E. (2003). *Methodological issues and strategies in clinical research* (3rd Edition). Washington, DC: APA.

Lee, J., & Bean, F. D. (2004). America's changing color lines: Immigration, race/ethnicity, and multiracial identification. *Annual Review of Sociology, 30,* 221-242.

Loving v. Virginia. No. 395. Supreme Ct. of the US. 12 June 1967.

Miville, M. L., Constantine, M. G., Baysden, M. F., & So-Lloyd, G. (2005). Chameleon changes: An exploration of racial identity themes of multiracial people. *Journal of Counseling Psychology, 52,* 507-516.

Morrison, J. W., & Bordere, T. (2001). Supporting biracial children's identity development. *Childhood Education, 77*, 134-137.

Morning A. (2003). New faces, old faces: Counting the multiracial population past and present. In H. DeBose & L. Winters (Eds.), *New Faces in a Changing America: Multiracial Identity in the 21st Century* (pp. 41-67). Thousand Oaks, CA: Sage.

National Association of Social Workers. (1999). *NASW Code of ethics*. Washington, DC: Author.

National Association of Social Workers. (2001). *NASW Standards for cultural competence in social work practice*. Washington, DC: Author.

Nishimura, N. (2004). Counseling biracial women: An intersection of multiculturalism and feminism. *Women & Therapy, 27,* 133-145.

Office of Management and Budget (1997). Office of Management and Budget: Revisions to the standards for the classification of federal data on race and ethnicity. Retrieved from http://www.whitehouse.gov/omb/fedreg/print/ombdir15.html.

Park, R. E. (1928). Human migration and the marginal man. *The American Journal of Sociology, 33,* 881-893.

Phinney, J. S. (1992). The multigroup ethnic identity measure: A new scale for use with diverse groups. *Journal of Adolescent Research, 7,* 156-176.

Poston, W. S. C. (1990). The biracial identity development model: A needed addition. *Journal of Counseling and Development, 69,* 152-155.

Radutzky, M., & Devine, L. F. (Producers). (2008, September 21). Obama '08 [Television series episode]. In D. Hewitt (Producer), *60 Minutes*. New York: CBS Broadcasting Company.

Renn, K. A. (2003). Understanding the identities of mixed-race college students through a developmental ecology lens. *Journal of College Student Development, 44*, 383-403.

Rockquemore, K. A., & Brunsma, D. L. (2002). Socially embedded identities: Theories, typologies, and processes of racial identity among black/bhite biracials. *The Sociological Quarterly, 43,* 335-356.

Rockquemore, K. A., Brunsma, D. L., & Delgado, D. J. (2009). Racing to theory or re-theorizing race? Understanding the struggle to build a multiracial identity theory. *Journal of Social Issues, 65*, 13 – 34.

Root, M. P. P. (1992). Back to the drawing board: methodological issues in research on multiracial people. In M. P. P. Root (Ed.), *Racially Mixed People in America* (pp. 181-189). Newbury Park, CA: Sage.

Root, M. P. P. (1992b). Within, between, and beyond race. In *Racially Mixed People in America* (pp. 3-11). Newburg, CA: Sage.

Root, M. P. P. (1998). Experiences and processes affecting racial identity development: Preliminary results from the biracial sibling project. *Cultural Diversity & Ethnic Minority Psychology, 4,* 237-247.

Root, M. P. P. (1999). Methodological issues in multiracial research. In G. C. N. Hall & S. Okazaki (Eds.), *Asian American psychology: The science of lives in context* (pp. 171-193). American Psychological Association.

Root, M. P. P. (1999b). The biracial baby boom: Understanding ecological constructions of racial identity in the twenty-first century. In R. H. Sheets & E. R. Hollins, (Eds.), *Aspects of Human Development: Racial and Ethnic Identity in School Practices* (pp. 67-90). Mahwah, NJ: Lawrence Erlbaum.

Root, M. P. P. (2001). On the development of the bill of rights for racially mixed people. In D. Goldberg & P. Essed (Eds.), *Race Critical Theories* (pp. 355-368). New York: Blackwell.

Root, M. P. P. (2003). Multiracial Families and Children: Implications for Educational Research and Practice. In J. A. Banks & C. A. McGee Banks (Eds.), *Handbook of Research on Multicultural Education* (2nd edition), pp. 110-124. San Francisco: Jossey-Bass.

Root, M. P. P., & Kelley, M. (2003). *Multiracial child resource book: Living complex identities.* Seattle, WA: MAVIN Foundation.

Samuels, G. M. (2006). Beyond the rainbow: Multiraciality in the 21st century. In D. Engstrom & L. Piedra (Eds.), *Our Diverse Society: Race, Ethnicity and Class—Implications for 21st Century America.* Washington, DC: NASW Press.

Samuels, G. M. (2009). "Being raised by white people": Navigating racial difference among adopted multiracial adults. *Journal of Marriage and Family, 71*, 80-94.

Shih, M., & Sanchez, D. T. (2005). Perspectives and research on the positive and negative implications of having multiple racial identities. *Psychological Bulletin, 131,* 569-591.

Shih, M., & Sanchez, D. T. (2009). When race becomes even more complex: Towards understanding multiracial identity and experiences. *Journal of Social Issues, 65*, 1-11.

Singh, G., & Johnson, M. (1998). Research with ethnic minority groups in health and social welfare. In C. Williams, H. Soydan, & M. R. D. Johnson (Eds.), *Social Work with Minorities: European Perspective* (pp. 231-246). London: Routledge.

Smith, J. P., & Edmonston, B. (1997). *The new Americans*. Washington, DC: National Academy Press.

Sommers, V. S. (1964). The impact of dual-cultural membership on identity. *Psychiatry, 27,* 332-344.

Stonequist, E. (1937). *The marginal man*. New York: Scribners.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory.* (2nd ed.) Thousand Oaks: SAGE.

Townsend, S. S. M., Markus, H. R., & Bergsieker, H. B. (2009). My choice, your categories: The denial of multiracial identities. *Journal of Social Issues, 65*, 183-202.

Udry, J. R., Li, R. M., & Hendrickson-Smith, J. (2003). Health and behavior risks of adolescents with mixed-race identity. *American Journal of Public Health, 93*, 1865–1870.

United States Bureau of the Census. (2000). *Current populations report*. Washington DC: U.S. Census Bureau.

U.S. Department of Health and Human Services. (2001). *Mental health: Culture, race, and ethnicity – A supplement to mental health: A report of the surgeon general.* Rockville, MD: U.S. Department of Health and Human Services, SAMHSA, Center for Mental Health Services.

Wardle, F. (1991). Interracial children and their families: How school social workers should respond. *Social Work in Education, 13,* 215-223.

Wilson, A. (1987). *Mixed race children: A study of identity*. London: Allen & Unwin.

Wijeyesinghe, C.L. (2001). Racial identity in multiracial people. In C. L. Wijeyesinghe & B.W. Jackson (Eds.), *New Perspectives on Racial Identity Development: A Theoretical and Practical Anthology* (pp.129-152). New York: University Press.

Finding Fathers in Social Work Research and Practice

Alyson F. Shapiro, Ph.D.
Arizona State University

Judy Krysik, Ph.D.
Arizona State University

Abstract

A comprehensive five-year review of five social work journals and one family-focused inter-disciplinary journal was conducted to examine the prevalence of recent research on fathers. Despite an increase in father-focused research over the past two decades, there continues to be a significant lack of research examining fathers relative to mothers, as well as research that includes fathers as participants. Ethical issues regarding the inclusion of fathers in social work research and practice are discussed.

Keywords: fathers, social work practice, family, research, ethics

1.0 Introduction

Research and social work practice with regard to families and parenting have traditionally focused on mothers, with fathers becoming increasingly recognized over the past three decades (Silverstein, 2002; Tamis-LeMonda & Cabrera, 1999). The lack of attention to fathers is inconsistent with social work practice frameworks, e.g., the ecological theory (Bronfenbrenner, 1979) and family systems perspectives (Minuchin, 1974), that are inclusive of all aspects of the family. Fatherhood became a topic of political focus with the Fatherhood Initiative in the 1990s (Tamis-LeMonda & Cabrera, 1999), and has recently become a widely recognized sub-specialty across a number of fields (see Doucet, Edwards, & Furstenberg, 2009). Despite this recent growth in father-focused research, the very conceptualization of fathering as a sub-specialty reflects that fathers are still not considered as central to parenting as mothers. There remains a great deal we do not know about fathers, particularly with respect to diversity and the effectiveness of father-focused interventions (Bayley, Wallace, & Choudhry, 2009; Doucet, Edwards, & Furstenberg, 2009). This is problematic for social workers, since it is difficult to engage and intervene with fathers without adequate knowledge.

The inclusion of fathers in social work practice and research is of ethical relevance since according to the National Association of Social Workers (NASW), "the primary mission of the social work profession is to enhance human well-being and help meet the basic human needs of all people…."(Preamble, NASW Code of Ethics, 2008). The NASW Code of Ethics was revised in 2008 to include several provisions facilitating the well-being of people without regard to sex or other aspects of cultural and social diversity. Specifically, "social workers should not practice, condone, facilitate, or collaborate with any form of discrimination on the basis of race, ethnicity, national origin, color, sex, sexual orientation, gender identity or expression, age, marital status…" (NASW section 4.02, 2008), and "social workers should obtain education about and seek to understand the nature of social diversity and oppression with respect to race, ethnicity, national origin, color, sex…." (NASW section 1.05 (c), 2008). These ethical guidelines both prohibit discrimination against fathers simply on the basis that they are the male parent, and highlight the importance of research on fathers that informs both social work practice and policy.

The focus of the current paper is not to provide a comprehensive review of the research on fathers, but rather to examine the recent inclusion of fathers in social work research, to review research

evaluating social work practice with fathers, and to discuss barriers related to practice with fathers. It is in the spirit of recognizing the need for equitable representation of both mothers and fathers in research and practice as highlighted by social work's dominant practice frameworks and the NASW Code of Ethics (2008) that this paper calls for the ethical consideration of the inclusion of fathers in family-focused research and practice.

2.0 Literature Review

As recently as 2002, Silverstein argued that due to the bias towards the maternal attachment paradigm, fathers continue to be a population we know little about, and that this is particularly true for fathers from diverse backgrounds including low-income, gay, and ethnic minority fathers. She called for research on fathers overall and qualitative research involving fathers in particular. Silverstein conceptualized the barrier to father inclusion in both research and practice in terms of gender theory, since nurturing is viewed primarily as feminine in U.S. culture. In a recent special issue of *Child Maltreatment* devoted to examining the father's role in child abuse and neglect, Lee, Belamy, and Guterman (2009) also highlighted the lack of research in this area, despite fathers being over-represented in the most severe cases of child maltreatment. Thus, striving towards a better understanding of fathers in nontraditional roles would be consistent with the NASW guidelines regarding cultural and social diversity (2008).

Levine and colleagues identified barriers to father involvement with regard to the social welfare community as well as contributing to the literature on how fathers can become increasingly involved across a number of settings such as early childhood centers and schools (Levine, Murphy, & Wilson, 1993; Levine & Pitt, 1995; Levine & Pittinsky, 1997). Social work tends to be a female dominated field, with many agencies being staffed primarily with women. Levine pointed out that many of these women have had negative experiences with men in their own lives that may make them reluctant to reach out to men. Furthermore, social workers and other mental health professionals are often not trained to work with families at the family-level, and to address the inter-parental conflict that is common when more than one parent is involved in dialogue regarding parenting issues. This evaluation of the climate of the social welfare culture is consistent with research indicating that both social work and psychology practitioners are more likely to include fathers in their practice if they are male (Lazar, Sagi, & Fraser, 1991).

As recently as 2006, Lee reported that fathers were rarely involved in psychological services related to their children's difficulties. This appears true for both resident and non-resident fathers (Duhig, Phares, & Birkeland, 2002). Phares, Fields, and Binitie (2006) identified a number of factors that may contribute to the lack of father participation including: therapists not actively inviting father participation, therapists' biases in not considering father participation important, discomfort with inter-parental conflict, fathers' time-constraints, fathers' assessment of therapeutic intervention as unnecessary, and fathers' problem solving or coping styles.

Research indicates, however, that increased father involvement is related to positive child well-being (Lamb & Tamis-Lemonda, 2004). Fathers tend to interact with their children in qualitatively different ways than mothers (see Lamb, 2004; Parke & Brott, 1999). Father rough and tumble play with positive affect predicts better child self-control abilities and peer acceptance (Snarey, 1993; Youngblade & Belsky, 1992), and father emotional involvement with pre-school aged children predicts later social competence (Gottman, 1997). Father involvement is related to children's school related success in both middle childhood (Nord, Brimhall, & West, 1997) and adolescence (Furstenberg & Harris, 1993). A combination of high father involvement and increased closeness appeared particularly important for buffering adolescents from distress and engaging in delinquent behavior (Harris, Fustenberg, & Marmer, 1998). Although some authors have described fathers as being essential (Pruett, 1998), and others regard fathers to be important but not essential (Silverstein, 2002), it is clear that fathers make important contributions to their children's lives.

In contrast, negative aspects of fathering and paternal psychopathology appear to be related to negative child outcomes. Father depression is related to child and adolescent internalizing and externalizing problem behaviors (Kane & Garber, 2004), and punitive parenting by the father is related to externalizing problems in male children (Heaven, Newbury, & Mak, 2004). Since positive parenting appears consistently associated with positive child outcomes, it would make sense for social workers to promote positive father involvement, and to intervene in cases where the father's parenting style appears punitive or harsh. This research also highlights the importance of identifying paternal as well as maternal mental health difficulties, and treating or referring these parents for treatment since their well being appears to have ramifications for the children as well as the parents.

Research evaluating social work practice with regard to fathers indicated father inclusion in permanency planning (Coakley, 2008) and parenting training (Lindahl, Tollefson, Risser, & Lovejoy, 2008) was associated with positive child outcomes, and couple-focused interventions appeared successful in promoting father involvement (Hawkins et al., 2008; Fagan, 2008). Excluding fathers from social work practice not only fails to improve child well-being through positive father involvement, but may also play an unintended role in minimizing father involvement. Research indicates that mothers play a gate-keeping role, moderating father involvement in families with both resident and non-resident fathers (McBride et al., 2005). Engaging mothers but not fathers in outreach and intervention further perpetuates the mother's gender ideology-based role as the primary caregiver, and may have the unintended consequence of excluding fathers.

Not only has the field of fatherhood research changed over the last several decades, there is evidence that the "culture of fatherhood" itself is also actively transforming (LaRossa, 1988). Fathers report that they want to be more involved with their children than their fathers were with them (see Parke & Brott, 1999), and indeed research indicates that contemporary fathers are more involved than those of previous generations (Pleck & Masciadrelli, 2004). Although fathers have traditionally been viewed as filling the role of the family breadwinner, they have more recently begun to be viewed as co-parents (Pleck & Pleck, 1997). Many fathers have started taking on longer child care shifts corresponding to an increase in maternal employment (Casper & O'Connell, 1998). This ongoing evolution of the face of fatherhood further necessitates continued research on fathers in order to ensure adequate understanding and engagement.

Diverse family compositions bring challenges to the definition of fathers and father figures with regard to both research and practice. Fathers can be biological, social (fictive kin), legal, or step-father, in families ranging from single-parent to married, cohabiting, and re-combined. Despite the difficulties involved in conceptualizing fathering across various family compositions, it is clear that father figures make contributions to their families across the spectrum. A national study of new births reported that even in single parent households, 75% of unmarried biological fathers had some level of ongoing involvement in the lives of their infant children, with 80% paying child support (Mincy, Garfinkel, & Nepomnyaschy, 2005). It is also noteworthy that as of 2004, 18% of single parent households in the U.S. were headed by fathers (U. S. Census Bureau, 2004). The active participation of non-resident dads has been related to increased child success in school (Nord, Brimhall, & West, 1997), and authoritative parenting style among non-resident dads appears positively related to school success and negatively related to child internalizing and externalizing problems (Amato & Gilbreth, 1999).

Most foster children are not living with their fathers when they are removed from their homes (Malm, Murray, & Green, 2006). Non-resident fathers could be an important resource for these children, particularly when these fathers are not implicated in the abuse. Malm and colleagues (2006) found that non-resident fathers are often not located early in the process of investigation, and that engaging with these fathers once located was rarely a stated priority since it was not legally mandated. CPS case workers who were not trained to engage fathers were also less likely to report that non-resident fathers had been identified. This lack of engagement with fathers was related to a decreased likelihood both that these children would be placed with their fathers, and that fathers would have some level of ongoing involvement in the lives of their children through visitation. Child welfare workers are encouraged to identify non-resident fathers early, and to engage with them to facilitate long-term father involvement.

In 1990, Grief and Bailey published the results of a comprehensive review of five major social work journals over a 27-year-period to examine the prevalence of research on fathers. They found only 21 articles focused on fathers, reflecting less than one father-focused article per year across the five journals examined. Furthermore, the articles on fathers focused narrowly on fathers as perpetrators, missing, and embattled. Their conclusion was that if social workers understood fathers, their understanding was not gleaned from their reading of major journals in their field (Grief & Bailey, 1990). How much has the parenting focus of social work journals changed since 1990? In an effort to examine the prevalence of recent research on fathers in social work and related fields, a five-year review of five widely recognized social work journals and one family focused inter-disciplinary journal was conducted. Articles evaluating social work practice with regard to fathers were identified and summarized to provide insight for discussion of ethical issues related to the involvement of fathers in social work practice.

3.0 Method

The social work journals examined were: *Social Work Research, Research on Social Work Practice, Children and Youth Services Review,*

Journal of Social Service Research, and *Social Service Review*. These journals were chosen due to their high citation ratings and clear relevance to social work. The one inter-disciplinary journal examined, *Family Relations*, was selected due to its high social work citation rating and specific focus on family related issues.

Titles and abstracts for each issue from 2004 through 2008 were first examined to identify the inclusion of family or parenting related variables. These articles were then further content analyzed to identify whether mother variables, father variables, or both, were included. If a study reportedly examined father variables whenever possible, for example in all dual-parent homes in a given sample, the article was considered to include a father variable. When research was reported for "caregivers," "parents," or "families," the breakdown of mothers, fathers and other caretakers was examined and recorded when possible. If an article included fathers, the source of the data regarding the father was further recorded to reflect: father inclusion in the study (father self-report or observation of father), and mother, caseworker, teacher or child report of father variables. Review articles were not considered in this examination. In addition to reporting descriptive statistics, t-tests were used to examine differences across groups.

4.0 Results

Across the six journals examined, 24% ($N = 62$) of the 262 family-focused articles included father variables, compared to 53% that examined mother variables, and 43% that generically examined "caregivers" or "parents." There is an overlap in the percentages presented due to most of the articles examining father variables also including mother variables. Specifically, only 2.6% of articles reported on father variables in the absence of mother variables, 21% included both mother and father variables, 31.7% included only mother variables, 42.7% reported on caregiver variables (not specifying a specific parent figure), and an additional 1.9% focused on grandmothers. This breakdown reflects significantly more articles examining mothers ($M = 4.86$) than fathers ($M = 2.86$; $t (29) = 5.28$, $p < .001$) per year, in an analysis where the N reflects the number of volumes examined across the six journals.

These differences were more distinct when examining only the five social work journals, which included 48% articles with mother variables, 48% with "caregiver" variables, and only 17% with father variables. This reflects more articles with mother variables ($M = 3.8$ per year) than father variables ($M = 1.28$ per year; $t (24) = 4.14$, $p < .001$). There was a significantly higher percent of articles that included

father variables in the inter-disciplinary *Family Relations* journal ($M = 34.68\%$) compared to the social work journals ($M = 15.05\%$; $t (28) = 4.88$, $p < .001$).

Investigation of the father variables examined revealed that only 54% employed either father report or observations of fathers. The remaining father variables reflected mother (12%), child (25%), case worker (7%) or teacher report (2%). Thus, although 24% of family-related articles examined fathers across the six journals reviewed, only 12.5% actually included fathers in the research. Within the five journals specific to social work, only 7.26% of the family-related articles included fathers in the research.

In articles where "caregiver" or "parent" variables were examined, a breakdown of the parent or caregiver figures was sought. However, in 43% of these articles, no clear breakdown was given, and thus it was not clear if these articles reflected exclusively mothers, or some combination of mothers, fathers, and other caregivers. The following statistics are reported for the articles where a breakdown of caregiver figures was given. The percentage of caregivers who were mothers or female ranged from 50% to 100%, with the average percent being 82.4%. Although it may be assumed that the remaining 17.6% of caregivers were fathers, not all of these articles gave specific breakdowns for fathers, and when breakdowns were given beyond the percent that were mothers, caregiver also often included grandmothers and foster parents.

Although most of the research examining father variables yielded information with implications for social work practice and policy, only nine of the articles clearly evaluated interventions or social work practice related to fathers. This breakdown does not include a handful of studies that included both mothers and fathers in intervention, but did not report results specific to fathers. A summary of the father evaluation research is presented below.

Davidson-Arad, Peled, and Leichtentritt (2008) examined court petitions for child removal written by child protection workers in Israel. The focus of these petitions was on blaming mothers and virtually ignored fathers. This pattern of mother blame and father absence was also the focus of a Canadian investigation of child protection case files (Strega et al., 2008). Strega and colleagues found that although a number of father-figures were identified, almost 50% of fathers were considered irrelevant to both child and mother. Father engagement by social workers was also low, with 60% of fathers considered a risk to children and not being contacted despite many having unsupervised visits with their children. In the face of this lack of engagement of

fathers by social workers, research by Coakley (2008) found that when African American fathers were involved in the permanency planning of their children in the child welfare system, more children had shorter stays in foster care, and were reunified with their birth families.

In another study, Davidson-Arad and colleagues attempted to predict social workers' decisions to remove children from families based on their early assessments of the children and parents (Davidson-Arad, Englechin-Segal, Wozner, & Arieli, 2006). This research found an interaction effect indicating that social workers rated fathers as more cooperative in families where the children were removed than both mothers in those same families and fathers in families where children remained placed in the homes. This result was not surprising given that mother, but not father, mental health difficulties were predictive of child removal. Thus, the mothers in those families may have been difficult to relate to because of their mental health difficulties, requiring the social workers to engage with the fathers. What is not clear from the report is what was going on in the families where the children remained in the homes. Were fathers truly less cooperative, or were social workers less likely to engage them to the same extent as mothers when the mothers were free of mental health difficulties?

Kohl and colleagues found that child welfare may be a gateway to domestic violence (DV) services (Kohl, Barth, Hazen, & Landsverk, 2005). Both identification of DV and having an open child welfare case predicted receipt of DV services. However, the focus of identification of DV and referral for related services was for the mothers only, and not their male partners.

Intervention research indicated that couple-focused (Hawkins et al., 2008) and co-parenting focused (Fagan, 2008) interventions can be successful in promoting father involvement. Finally, an evaluation of a family nurturing program for incarcerated and at-risk participants indicated that men demonstrated increased understanding of developmentally appropriate expectations, empathy, and the impact of corporal punishment (Palusci, Crum, Bliss, & Bolovek, 2008). These combined results reflect the positive potential for including fathers in intervention.

5.0 Discussion

Research on fathers published in social work journals has clearly proliferated, with 62 articles including father variables being evident over the past five years compared to only 21 reported in Grief and Bailey's 1990 review of a 27-year-period. There

continues to be, however, a significant lack of research including fathers relative to mothers in family-related research, with only 24% of family-focused articles including father variables across the six journals examined, and only 12.5% including fathers as participants in the actual research. These results are even more significant for the journals examined specific to social work, with only 17% of family-focused articles including father variables, and only seven percent actually including fathers in the research. This relative lack of father-related research is problematic for social work, since it is this research that is most likely to inform practice and policy. The finding that conclusions regarding fathering in these journals largely come from sources other than the father is also problematic, since these findings could actually be perpetuating the gender biases of female parents and practitioners, resulting in continued inequity in the social welfare system with regard to fathers. This is of particular concern from an ethical standpoint, since it could interfere with the ability of social work practitioners to adequately understand and address the needs of the fathers and the children in the families they serve. Issues prohibiting discrimination against people with regard to sex and advocating the seeking of adequate knowledge across sexes are explicitly addressed in the NASW Code of Ethics (2008).

Why is there only a small percent of articles actually including fathers in the research despite a clear increase in awareness of the importance of examining fathers? One explanation is that a number of the research articles reported secondary analysis of data available from large national studies or administrative databases. In these cases, often only mothers and their families were examined, and only mother report of father variables was available. However, mother and practitioner report of father related issues should not be considered synonymous with father observation or report given research indicating that there can be low concordance between mother and father report even on issues that may overtly seem straightforward, such as report of the father's occupation (Schnitzer, Olshan, Savitz, & Erikson, 1995).

Researchers are encouraged to include fathers as well as mothers in their research designs to the extent possible. Furthermore, granting agencies and grant reviewers are encouraged to consider the inclusion of fathers as well as mothers in the funding of research, particularly when reviewing large multi-site research studies that may be the basis of secondary as well as primary data analyses over the decades to come. Several articles reviewed reported on data from the National Longitudinal Survey of Youth (see for example Grogan-Kaylor & Otis,

2006). Unfortunately, this study followed only the families of adolescents who became mothers and did not include the adolescents who became fathers. In contrast, a number of the articles reporting research including fathers came from secondary data analyses of the Fragile Families and Child Wellbeing Study (see Huang & Warner, 2005). More longitudinal, multi-site studies with rich data being gathered from both mothers and fathers are needed to facilitate further proliferation of research that actually includes fathers in the research.

A sizable 42% of articles with a family focus reported data for caregivers or parents rather than specifically gathering information on mothers or fathers. Although at first glance this may seem like an egalitarian approach, there is evidence suggesting that disparate caregiver figures should not be lumped into one category. Research indicates that fathers make contributions to their children's development independent of mothers (Tamis-LeMonda & Cabrera, 1999), and that fathers and mothers tend to provide unique relational settings for child development (Ryan, Martin, & Brooks-Gunn, 2006). Thus, it would be a mistake to view different caregiver figures as synonymous, and this view is likely to carry over into the area of social work practice. In cases where caregivers are examined that include different parent figures, it is critical to include a clear breakdown by gender. Furthermore, researchers are encouraged to examine differences between types of caregivers in cases where sub-groups of caregivers, such as fathers and mothers, are large enough to allow such statistical comparisons. The changing culture of fatherhood also has implications for social work practice. It is one thing to work primarily with the mother in a context where both parents are content with the role of the mother as the primary parent figure, and another more ethically problematic thing to focus primarily on mothers in a context where fathers want to be more involved. Thus, as fathers become increasingly involved with their children, and express the desire to do so, it becomes increasingly important for social workers to make efforts to include fathers in their family-related practice. Overall, intervention research indicates that including fathers in social work practice has the potential to enrich the lives of their children when and if social workers identify, locate, and engage with these father figures.

The relative deficiency of father inclusion in both social work practice and family focused research is also likely to be due to both real and perceived barriers to father participation that are even greater when the father resides apart from the mother and child. (Bayley, Wallace, & Choudhry, 2009; Lee, 2006). Logistical arrangements that facilitate the inclusion of fathers as well as mothers includes actively soliciting father participation and conducting therapeutic appointments and home visits during flexible times, often during evenings and weekends to accommodate the scheduling needs of one or more working parent. Although fathers may not agree to be involved in psychological services as readily as mothers (Duhig, Phares, & Birkeland, 2002), when active efforts are made to include fathers and accommodate their scheduling needs, research indicates they participate to some extent in both intervention and research (Phares, 1996; Phares, Fields, & Binitie, 2006; Shapiro & Gottman, 2005).

Couple conflict regarding parenting issues, along with a wide array of other issues, is not uncommon. Thus, it is not surprising that Levine and colleagues indicated that when fathers are included in social work practice, couple conflict often ensues (Levine, Murphy, & Wilson, 1993; Levine & Pitt, 1995; Levine & Pittinsky, 1997). This couple conflict can result in social workers and other practitioners experiencing feelings of discomfort, and they may even be concerned that they have contributed to an escalation of negativity in the family by including fathers as well as mothers in their outreach. This in turn can lead to practitioners reverting to working primarily, if not exclusively, with mothers. It is important to recognize that inter-parental conflict is an integral part of couple relations and is likely to exist whether or not practitioners are exposed to it. Inter-parental conflict can be viewed as an opportunity for exploration of underlying views and for building conflict resolution skills. Referring families for marital or couple counseling may be appropriate in some cases, but practitioners are cautioned against avoiding the inclusion of fathers in their practice simply due to parental conflict. For parents who have engaged in domestic violence or child abuse or neglect, the tendency may be exclusion, however, there is support that these men may also benefit from an increased understanding of the effects of violence on their developing children. These considerations have implications for the education of social workers beyond what is currently offered.

References

Amato, P. R., & Gilbreth, J. G. (1999). Nonresident fathers and children's well-being: A meta-analysis. *Journal of Marriage and the Family, 61,* 557-573.

Bayley, L. K., Wallace, L. M., & Choudhry, K. (2009). Fathers and parenting programs: Barriers and best practice. *Journal of Community Practice, 82*, 28-31.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Cambridge, MA: Harvard University Press.

Casper, L. M., & O'Connell, M. (1998). Work, income, the economy, and married fathers as child care providers. *Demography, 35*, 245-250.

Coakley, T. M. (2008). Examining African American fathers' involvement in permanency planning: An effort to reduce racial disproportionally in the child welfare system. *Children and Youth Services Review, 30,* 407-417.

Davidson-Arad, B., Englechin-Segal, D., Wozner, Y., & Arieli, R. (2005). Social workers' decisions on removal: Predictions from their initial perceptions of the child's features, parents' features, and child's quality of life. *Journal of Social Service Research, 31,* 1-24.

Davidson-Arad, B., Peled, E., & Leichtentritt, R. D. (2008). Representation of fathers and mothers in court petitions for dependent minor status for children at risk. *Children and Youth Services Review, 30,* 893-902.

Doucet, A., Edwards, R., & Furstenberg, F. F. (2009). *Fathering across diversity and adversity: International perspectives and policy interventions.* Thousand Oaks, CA: Sage.

Duhig, A. M., Phares, V., & Birkeland, R. W. (2002). Involvement of fathers in therapy: A survey of clinicians. *Professional Psychology: Research and Practice, 4,* 389-395.

Fagan, J. (2008). Randomized study of a prebirth coparenting intervention with adolescent and young fathers. *Family Relations, 57,* 309-323.

Furstenberg, F., & Harris, K. (1993). When and why fathers matter: Impacts of father involvement on the children of adolescent mothers. In R. Lerman & T. Ooms (Eds.), *Young Unwed Fathers: Changing Roles and Emerging Policies* (pp. 117 – 138). Philadelphia: Temple University Press.

Gottman, J. M. (1997). *The heart of parenting.* New York: Simon & Schuster.

Grief, G. L., & Bailey, C. (1990). Where are fathers in social work literature? *Families in Society, 72,* 88-92.

Grogan-Kaylor, A., & Otis, M. D. (2006). The predictors of parental use of corporal punishment. *Family Relations, 55,* 80-91.

Harris, K., Furstenberg, F. F., Jr., & Marmer, J. K. (1998). Paternal involvement with adolescents in intact families: The influences of fathers over the life course. *Demography, 35,* 201-216.

Hawkins, A. J., Lovejoy, K. R., Holmes, E. K., Blanchard, V. L., & Fawcett, E. (2008). Increasing fathers' involvement in child care with a couple-focused intervention during the transition to parenthood. *Family Relations, 57,* 49-59.

Heaven, P. C., Newbury, K., & Mak, A. (2004). The impact of adolescent and parent characteristics of adolescent levels of delinquency and depression. *Personality and Individual Differences, 36,* 173-185.

Huang, C., & Warner, L. A. (2005). Relationship characteristics and depression among fathers with newborns. *Social Service Review, 79,* 95-118.

Kane, P., & Garber, J. (2004). The relations among depression in fathers, children's psychopathology, and father-child conflict: A meta-analysis. *Clinical Psychology Review, 24,* 339-360.

Kohl, P. L., Barth, R. P., Hazen, A. L., & Landsverk, J. A. (2005). Child welfare as a gateway to domestic violence services. *Children and Youth Services Review, 27,* 1203-1221.

Lamb, M. E. (2004). *The role of the father in child development*, 4th ed. New York: Wiley.

Lamb, M. E., & Tamis-Lemonda, C. S. (2004). The role of the father: An introduction. In M. E. Lamb (Ed.), *The Role of the Father in Child Development* (pp. 1 – 31), 4th ed. New York: Wiley & Sons.

LaRossa, R. (1988). Fatherhood and social change. *Family Relations, 36,* 451-458.

Lazar, A., Sagi, A., & Fraser, M. W. (1991). Involving fathers in social services. *Children and Youth Services Review, 13,* 287-300.

Lee, L. M. (2006). Involving fathers in the delivery of psychological services. *Cognitive and Behavioral Practice, 13,* 40-41.

Lee, S. J., Bellamy, J. L., & Guterman, N. B. (2009). Fathers, physical child abuse, and neglect: Advancing the knowledge base. *Child Maltreatment, 14,* 227-231.

Levine, J. A., Murphy, D. T., & Wilson, S. (1993). Getting men involved: Strategies for early childhood programs. New York: Scholastic.

Levine, J. A., & Pitt, E. W. (1995). New expectations: Community strategies for responsible fatherhood. New York: Families and Work Institute.

Levine, J. A., & Pittinsky, T. L. (1997). *Working fathers: New strategies for balancing work and family.* New York: Addison-Wesley.

Lindahl, B. W., Tollefson, D., Risser, H., & Lovejoy, M. C. (2008). A meta-analysis of father involvement in parent training. *Research in Social Work Practice, 18(2)*, 97–106.

Malm K., Murray J., & Green, R. (2006). *What about the dads? Child welfare agencies' efforts to identify, locate and involve nonresident fathers.* Retrieved from http://www.urban.org/publications/411316.html.

McBride, B. A., Brown, G. L., Bost, K. K., Shin, N., Vaughn, B., & Korth, B. (2005). Paternal identity, maternal gatekeeping, and father involvement. *Family Relations, 54,* 360-372.

Mincy, R., Garfinkel, I., & Nepomnyaschy, L. (2005). In-hospital paternity establishment and father involvement in fragile families. *Journal of Marriage and the Family, 67,* 611-627.

Minuchin, S. (1974). *Families and family therapy.* Cambridge, MA: Harvard University Press.

National Association of Social Workers (2008). National Association of Social Workers *Code of Ethics.* Approved 1996 and revised 2008 by the NASW delegate assembly. Retrieved from http://www.socialworkers.org/pubs/code/code.asp.

Nord, C. W., Brimhall, D., & West, U. (1997). *Fathers' involvement in their children's schools.* Washington, DC: U.S. Department of Education. ED 409 125.

Palusci, V. J., Crum, P., Bliss, S. J., & Bolovek, S. J. (2008). Changes in parenting attitudes and knowledge among inmates and other at-risk populations after a family nurturing program. *Children and Youth Services Review, 30,* 78-89.

Parke, R. D., & Brott, A. A. (1999). *Throwaway dads: The myths and barriers that keep men from being the fathers they want to be.* New York: Houghton Mifflin Company.

Phares, V. (1996). *Fathers and developmental psychopathology*. New York: Wiley.

Phares, V., Fields, F., & Binitie, I. (2006). Getting fathers involved in child-related therapy. *Cognitive and Behavioral Practice, 13,* 42-52.

Pleck, J. H., & Masciadrelli, B. P. (2004). Parental involvement by U.S. residential fathers: Levels, sources, and consequences. In M. E. Lamb (Ed.), *The Role of the Father in Child Development* (4th ed., 222-271). New York: Wiley.

Pleck, E. H., & Pleck, J. H. (1997). Fatherhood ideals in the United States: Historical dimensions.

In M. E. Lamb (Ed.), *The Role of the Father in Child Development* (3rd ed., pp. 33-48). New York: Wiley.

Pruett, K. D. (1998). How men and children affect each other's development. *Zero to Three, 18,* 3-11.

Ryan, R. R., Martin, A., & Brooks-Gunn, J. (2006). Is one good parent good enough? Patterns of mother and father parenting and child cognitive outcomes at 24 and 36 months. *Parenting: Science and Practice, 6,* 211-228.

Schnitzer, P., G., Olshan, A. F., Savitz, D. A., & Erickson, J. (1995). Validity of mother's report of father's occupation in a study of paternal occupation and congenital malformations. *American Journal of Epidemiology, 141,* 872 – 877.

Shapiro, A. F., & Gottman, J. M. (2005). Effects on marriage of a psycho-communicative-educational intervention with couples undergoing the transition to parenthood, evaluation at 1-year post intervention. *Journal of Family Communication, 5,* 1-24.

Silverstein, L. B. (2002). Fathers and families. In J. P. McHale & W. S. Grolnick (Eds.), *Retrospect and Prospect in the Psychological Study of Families* (pp. 35-64). New Jersey: Lawrence Erlbaum Associates.

Snarey, J. (1993). *How fathers care for the next generation.* Cambridge: Harvard University Press.

Strega, S., Fleet, C., Brown, L., Dominelli, L., Callahan, M., & Walmsley, C. (2008). Connecting father absence and mother blame in child welfare policies and practice. *Children and Youth Services Review, 30,* 705-716.

Tamis-LeMonda, C. S., & Cabrera, N. (1999). Perspectives on father involvement: Research and policy. *Social Policy Report: Society for Research in Child Development, 13,* 1-16.

U. S. Census Bureau. (2004). *America's families and living arrangements: 2004.* http://www.census.gov/population/www/socdemo/hh-fam/cps2004.html.

Youngblade, L. M., & Belsky, J. (1992). Parent-child antecedents of 5-year-olds' close friendships: A longitudinal analysis. *Developmental Psychology, 28,* 700-713.

Client-Centered Evaluation: Ethics for 21st Century Practitioners

Martin Bloom, PhD
Professor Emeritus
University of Connecticut

Abstract

Ultimately, evaluation belongs in the hands of the person most affected, the client, who should testify whether a conventionally evaluated intervention has attained his or her desired goal, as well as demonstrate that he or she can perform this intervention independently. Conventional evaluation is simultaneously performed on behalf of society to document the effective and efficient performance of public-financed service.

Key words: Client-centered evaluation, ethics, intervention phase, maintenance phase, single-system designs

1.0 Ethics in Evaluation and Research

Ethics, briefly defined with reference to the helping professions, involves a set of principles of right conduct that is supposed to govern practitioners' behaviors in clinical and social change situations (Reamer, 2006; Reamer & Shardlow, 2009; NASW Code of Ethics, 1999). There is no "ethics army" so that the enforcement of its principles is usually left to the consciences of mature individuals in the professions. As a constant reminder, these professions usually have concretized their ethical principles in written codes that are derived, more or less directly, from Hippocrates (circa 460 -377 B.C.E.). Veach (1981:22) quotes Hippocrates' The Epidemics as follows: "As to diseases, make a habit of two things – to help, or at least to do no harm." These 18 words or their equivalent have survived for nearly twenty-five hundred years to guide medical and other helping practices (Hartsell, 2006;

Loewenberg, Dolgoff, & Harrington, 2000; Meacham, 2007; Reamer, 2006). It is the task of this paper to connect 21st century evaluation with 5th century B.C.E. ethics.

I make the distinction between research and evaluation because the ethical requirements and implications of each differ in some important ways. This is true in spite of other applications of the terms, such as evaluative research, program evaluation, and the like, which constantly blend the two concepts as if they were the same thing. *Evaluation*, in the narrow sense in which I am using the term, involves producing practical and approximate knowledge for immediate use in client situations to provide guidance for achieving client goals (Bloom, Fischer, & Orme, 2009). *Research* involves producing abstract but usually mathematically specific knowledge for long-term understanding of conditions surrounding a group of persons with common problems or concerns. Research rarely provides any feedback to the subjects who composed the study group. Research studies supply the evidence for evidence-based general practice; my focus here is on evaluation-informed specific practice, that is, once a general intervention is selected based on the best available research, then practitioners have to individualize it or customize it in their intervention plans and then monitor its effectiveness with a specific client.

Evaluation involves the comparison of one client-system's current situation against some other reference point such as comparing back to that system's own initial baseline starting point (for example, what I weighed two months ago compared to what I weigh now); or comparing forward to some external standard or benchmark (like my trying to lose weight so as to reach the weight limits required

for joining the police force). Thus, evaluation uses the client as his or her own "control" group, a perfect equivalence with which to make these "before/after" comparisons, an equivalence that classical research designs and randomization seek to emulate – but never fully succeed in achieving. Evaluation, unlike research, is intentionally constructed to shape practice in field settings at the moment information is needed to make practical decisions (such as proceed as planned, change out of a deteriorating situation, or terminate as having achieved a stable goal). Classical research, often called the "gold standard" for empirical information, usually requires much more time for funding, institutional reviews, completion, and analysis, let alone the rare translation into concrete practice suggestions for clients/subjects in general, so that this gold is seldom transmuted into everyday currency to shape practice in real time.

Evaluation, in contrast to research, is very much influenced by its participants, which includes on-the-dime changes of direction when local evidence supports such change; additions of new targets as needed; and consideration of client's specific positives (to be maintained) as well as that client's problems (to be resolved) – any of which would not fit well into classical research designs, which more likely deal with problems in general. Evaluation is sensitive to the nuances of the particular client situation, in distinction to classical research, which has to consider the same outcome measures for all participants. The approach evaluators take has enormous advantages in the immediate client situation. For example, on-going measurement and monitoring makes it possible to detect signs of deterioration early and thus be able to make suitable adaptations. Of course, the same monitoring can inform practitioners about early signs of probable success, which allows planning for additional interventions if needed. The point is that evaluation is time-focused on the here-and-now, while research holds a more futuristic view of outcomes, namely, those that are the basis for evidence-based general practice.

These and like considerations have led my colleagues and me to consider a form of evaluation we call single-system design (Bloom, Fischer, & Orme, 2009) because of its applicability to a wide array of immediate practice situations involving persons, groups, or collectives. The chief characteristics of this approach may be briefly described as follows: single-system designs involve a set of empirical procedures to observe changes in an identified target (the dependent variable, a problem or objective of the client) that is measured repeatedly over time using the same procedures. A baseline reference pattern is used to compare the same

targeted pattern during intervention with reference to client goals or other benchmarks for desired outcomes. Some of the more elaborate designs used in single-system evaluation permit inferences of causality, such as multiple baseline designs or experimental repetition designs (A-B-A-B). The basic (or minimal scientific) A-B design compares baseline against intervention, which logically permits objective assessment of change, without specifying causal factors.

I call attention to the fact that at times in most practitioners' lives, they need to know the causal situation, especially in the maintenance phase when they teach the client to use a successful intervention on his/her own, as I will discuss shortly. These reasons are very different from why a researcher wants to have causal information, to build the evidence basis of social science. In conventional single-system evaluations, data are analyzed visually and/or by statistical methods, and practical decisions are made using the results, in conjunction with other considerations (such as agency policy, values, and professional standards). It is this immediate practicality of evaluation results that is critical, compared to the longer term research results that eventually reach a published form, which becomes the basis of evidence-based general practice.

There are threats to the validity of the findings in A-B designs that more elaborate evaluation designs can minimize, but all evaluations are tentative, approximate measures of a changing state of affairs that are to be used within the entire context of information, something like watching a moving picture of the client's targeted life. Thus, the practitioner cannot say with absolute certainty that evaluation results are the final word on outcomes; rather, these immediate outcomes themselves have to be assessed by the client as exhibiting a positive and meaningful change in that client's life. Such personalized results are in contrast to the usual statistical results that summarize or express changes in research for the population studied.

2.0 Client-Centered Evaluation in Practice

I use the term client-centered evaluation to represent a new configuration of choices and decisions practitioners and clients have to make together. Let me enumerate them:

First, a client-centered evaluation involves the client as much as possible in identifying the *goals* of intervention. The phrase, client-desired outcomes, names the process by which goals of the case are first enunciated. The practitioner usually has to re-state these desired outcomes to ensure that they are

publicly clear and agreed on, so that the rest of the intervention can focus on attaining them.

Second, goals are broken down into feasible intermediary steps, called *objectives.* Targets are the operational proxies for objectives, in which the specific client concern is selected for intervention by means of a specific set of actions, as clarified by the practitioner. The measured degree of a given target is plotted on a graph, and over time, these data represent a moving image of the target. Often clients are involved in collecting data as well as in discussions about changes in targets.

Third, as usual, *baselines and intervention phases* are conducted as needed to provide a logical basis for observed changes in the target as in the basic AB design. Advanced designs permit the logical inference of causality, such as ABAB and multiple baseline designs.

Fourth, regardless of any improvement in client functioning that occurs in the B phase of an AB design or the second B of the ABAB design, I would assert that it is unethical to stop here. Rather, I propose that the basic evaluation design be described as an ABM design, where M refers to the *maintenance phase* in which the client is instructed on how to take over the entire machinery of his or her improvement (to the extent possible) and demonstrate that he or she can continue obtaining positive results without practitioner assistance for a sustained time period. This time period will vary depending on the seriousness of the concern, but stability in desired outcomes is a basic condition for client-centered practice. Then, and only then, will the client's goals of attaining desired and sustained outcomes be a reality, so far as the practitioner can make this happen. Moreover, this dual analysis – a logical outcome in a B phase and the psychological outcome of the M phase – emphasize the dual ethical concerns for societal welfare and personal well-being reflected in the current NASW Mission Statement. The same would be true of any ABABM design, and in a multiple baseline design in which the client has to demonstrate independent control over any client-confirmed positive outcome.

3.0 Hippocrates in the 21st Century

The great possibility of twenty-first century client-centered evaluation's contribution to ethics is to fulfill its goals. Say you want to *help* if you can? Good, practitioners now possess the tools for each client in almost every clinical or social setting to evaluate whether change has occurred, and possibly whether they have contributed to that change. But is it no longer acceptable to say merely that "we helped." Ethical accountability demands that

practitioners specify what was achieved with the help that was given. More than that, they must specify to what degree they achieved the goals that were sought by clients and society. And further, they must indicate at what costs. These issues will lead us into an ancient ethical discussion, in a few moments.

Practitioners also need to deal with *do no harm* by returning to some ancient ethical issues that play out dramatically in our own day. Ultimately, I would argue that we can never say that "no harm has been done" because helping an individual client is like every surgical procedure that "harms" or cuts into the integral body surface in order to bring about a greater healing. Ethics is a comparative enterprise. Practitioners have to weigh what it costs to achieve what desired effects, and with what costs to the parties involved. There is no free lunch.

Likewise, evaluators, unlike researchers, must be aware at all times that it is never easy to be a client seeking help, effectively admitting failure at self-resolution or local solutions by family and friends. That cuts deeply into the integral personality, but it is a necessary "harm" to get the process of healing started. By engaging both the strengths of clients, as well as their presenting concerns and problems, we actively minimize harms. Research does not concern itself with such nuances, especially with the control group during the research period. This is demonstrated in the mathematics of research where all subjects are effectively considered equivalent, which practitioners know is never the case.

4.0 Evaluation in Ethical Theories, Ancient and Modern

Ethicists tell us that there are two broad classes of ethical theory that can be described as 1) the absolutist or deontological approach, and 2) the consequentialist or teleological approach (Meacham, 2007). There are many positions in between, taking elements of each model and combining them, often in Rube Goldberg fashion, to generate a new theory of ethics. The absolutist approach blends with the all-or-nothing position, which asserts that either we helped achieved client goals, or we didn't. Our duty is to reach this level of perfection and to keep matters there. Then the client and we live happily ever after.

I find the absolutist approach to be fairy-tale-like, and unsuited to the real world. The consequentialist approach tells us that by the fruits will you know whether the objectives have been achieved, especially if you can ask the client at risk. Most importantly, this position recognizes that everything under the social/cultural sun comes in degrees, more of this, less of that. This is true of

evaluation as well, which leads to the obvious question: how much of a desired outcome has to be achieved before we accept the intervention as producing "help (if you can), or at least do no harm"?

The issue, namely, what do we set as standards of successful outcome, is complex. Let me raise a few questions: First, recognize that practitioners can obtain degrees of improvement (or deterioration) compared to a baseline pattern. But how different must these two patterns (baseline and intervention phase data) be before victory is declared and we all go home? Let's assume we have a clear operationally defined target, such as minutes Junior practices each day on the piano, or pounds pudgy Sister loses after a period of vigorous exercise and controlled eating, or the degree to which aged Mr. Smith feels less angry for having been effectively forced to move into a nursing home after he lost practically every social support he had. What is measurably good/desirable/moral in these cases?

Perhaps the music teacher, based on years of experience, sets the goal of 30 minutes of practice a day (no procrastination, real playing). Maybe the school nurse has suggested to Sister's parents that she exercise X minutes and eat Y calories a day, which has been shown to take off pounds. Maybe there is no goal for Mr. Smith, other than survival at some acceptable level of contentment.

Evaluation of the first of these examples (piano playing) is simple, and both child and parent can verify the number of real practice time to some acceptable level of reliability. The second example is more complex because both exercise and eating have to be monitored, which can be difficult when Sister is on her own (and hungry) at school. The third example is very difficult, since there are no norms for contentment under these conditions, and Mr. Smith himself might doubt that nursing home living is really living at all.

This leads to a second consideration on ethical judgments when no clear goals or norms are available. In these instances, we are likely to fall back on science as being neutral and disinterested (compared to practitioners who are not neutral and are definitely interested in positive outcomes). Are the patterns at intervention significantly different than those at baseline?
But notice that this question, frequently asked in the literature on single-system design, will not tell us if Mr. Smith is contented. It only tells us that his contentment level (however we choose to measure it) is significantly different between the two time periods. And even this is relative: if Mr. Smith was at the 5% level of contentment at first, and then zoomed up to 10% after this intervention, even 10% may not be of any practical significance in Mr. Smith's life.

Say we have norms on contentment from a variety of studies of older people who had been moved into nursing homes from various earlier living sites. Let's say 40% of these elders were contented with their new surroundings. Or, let's say 4% were contented. The point is that for any pattern of change comparing Mr. Smith's scores with these general findings, we can make statistical interpretations, such as that Mr. Smith's 10% level of contentment is far below the 40% norm, or slightly above the 4% norm, which is an important first step in evaluation.

However, this is only the first step in an ethical analysis. I believe that ultimately, the client has to interpret any change within that client's own perspective. Mr. Smith could say that his level of contentment (at 10%) is not satisfactory, no matter whether it is compared to the 40% or the 4% norm. Unless we build this client determination directly into our evaluation process, we are avoiding the ultimate ethical decision regarding this client's situation.. This is not to say that I have any less enthusiasm for repeated systematic observations or the rest of the machinery of single-system designs, but rather I have come to appreciate more fully the operationalized fulfillment of client ethical concerns.

5.0. Research Evidence Versus Evaluation Evidence

Let's take a second tack in looking at research and evaluation in the ethical context. The rising tide of science in support of practice has been labeled evidence-based practice, and I do not intend to surf against the tide. However, with single-system designs, the practitioner has a marvelous surf board to ride with the wave.

Let's begin by defining terms, since evidence-based practice comes with many associated concepts (empirically-based practice, scientific practice, among them). "Evidence-based practice represents the practitioner's commitment to use all means possible to locate the best (most effective) evidence for any given problem at all points of planning and contact with clients" (Bloom, Fischer, & Orme, 2009, p.13). However, the underlying meaning of this definition includes one more term: evidence-based *general* practice. This means that the analysis and combination of many studies leads to general conclusions for practice, such as the probability that such-and-such a procedure, if properly executed, will likely lead to this pattern of results. This is vital information, and propels the tide in favor of universal evidence-based practice. And this is why we use the enlarged phrase, evidence-based general practice (Bloom, Fischer, & Orme, 2009).

However, what practitioners need, after they have some evidence-based general information on what works under what conditions for what kinds of people and problems, is much more specific: what will work with my client sitting in front of me? This is where evaluation enters the professional scene. Having selected one (or more) methods from the evidence provided by analysis of studies on this clinical or social change topic, the practitioner has to tailor the general methods, questions, time table, to fit the conditions of the immediate client. Recall that NASW Code of Ethics requires social workers to base practice on recognized knowledge, including empirically based knowledge (4.01.c). How will they know if this general evidence is working, especially having made some modifications to suit the conditions for their specific clients? By evaluating their own practice.

Using the evidence-based general practice ideas, the practitioner now fine tunes his or her evaluation to become informed about the specific effects with a given client. Hence, we call this evaluation-informed specific practice (Bloom, Fischer, & Orme, 2009). I now want to connect this to ethical considerations: By combining both evidence-based general practice and evaluation-informed specific practice, and involving clients (as far as possible) in choice of goals, demonstrating control over a successful intervention, and determining when an accomplished outcome in fact attains those goals, we have come one step closer to ethical practice for the 21st century. To paraphrase Kant, evidence-based general practice without specific evaluation is clumsy; evaluation-informed specific practice without general evidence is blind. Both are necessary; both must be taught to the new generations of ethical helping professionals.

6.0 Specific Ethical Considerations in Using Single-System Evaluation

I now turn to some of the specifics in using ethical methods of evaluation within the single-system design paradigm. Let's begin with the fundamental assumption, that any and every choice and action a practitioner makes in the field of practice with clients is an ethical decision (Reamer, 2006). Such choices and actions presumably introduce changes to the client's life -- that is to say, in his or her interactions with others, which is the moral ground of all human behavior. Here are the ethical particulars in single-system evaluation:

1. We must seek to provide *demonstrable help* for the individual client in his or her social setting. For this, a single-system design provides the closest form of hard evidence that we can attain

easily, quickly, at low costs, without the use of elaborate methods. Evaluation is an approximation of the best research; sometimes the approximation is very close (with advanced single-system designs); other times, evaluation provides only hints (with the basic A-B design).

2. We must demonstrate that *no harm* was done in the process of intervening and collecting ongoing data, for both the individual client and for his or her social context. Thus, even if help could be provided for the client while at some harm to others in the social context, the principle of doing no significant harm within the larger social context takes precedence over doing good for the client (Reamer, 2006; but also see Hartsell, 2006). The single-system form of evaluation can make these distinctions, which would likely be hidden by group data in research -- and likewise in using only evidence-based general research for our unique client. Single-system design has a significant contribution to the dual focus of combining evidence-based general practice with evaluation-informed specific practice.

3. Since ethical considerations involve the client directly, so in evaluation, we often involve the client as directly as possible in selecting targets, choosing methods of data collecting, participating in those data collections, and interpreting the outcomes as part of real world changes the client is seeking. (We may not involve some clients directly, if they have limitations of age or functioning, in which case their adult caretakers are directly involved.)

4. Evaluation has the luxury of stopping itself, should the intervention or data collection prove to be painful or harmful to the client or client-system, physically, psychologically, or socially. The difficult part is performing this action without prejudicing the services being offered. The clinical and lay literature is littered with terms such as "noncompliant patient" (Bloom, 2008) and the like -- and I have no doubt that clients can be difficult -- but in fact clients are telling us something by "acting out" or "violating agreements" that we don't like to hear, but must hear if we are to act appropriately. Evaluation offers many ways to achieve its results that may not involve actions clients find offensive, like having to report on their own behavior, by means of indirect methods including unobtrusive observer ratings made in private. In research, we are generally stuck with the methods and designs we started with.

5. Evaluation shares with research the obligation of confidentiality with regard to data, records, and the like. However, it may be easier for the one-on-one practitioner/client situation to explain the limits of confidentiality (see Tarasoff v. Regents of California [17 Cal. 3rd 425, 1976; Kopels & Kagle, 1993). The difference between confidentiality

in evaluation as contrasted with research is that a bond of trust is likely to exist between the practitioner and client in the former that does not exist in the latter, that all information obtained in confidence must benefit all parties and harm none. It should be obtained through informed consent (Bloom, Fischer, & Orme, 2009: 572).

6. Evaluation takes into consideration directly the unique factors of the individual client, including ethnicity, income level, educational level, sexual orientation, and gender. These become factors to be controlled in research, and yet in evaluation, they are the specific guiding factors on which decisions have to be made. It was for this reason that we constructed a client bill of rights (Bloom, Fischer, & Orme, 2009: 571) that explicitly states what are client rights regardless of individual differences (Wilson, 1983). With these kinds of sensitivities in evaluation, we can increasingly know where this

specific client is (regarding targets); we can reduce the drop out rate (which often occurs when the client is dissatisfied with the intervention); we can have fewer treatment sessions with no worsening of outcomes for the clients making progress -- this will be shown in increased cost-effectiveness (Harmon, et al., 2007; Lambert, 2007).

7.0 Summary: Client-Centered Evaluation in Practice

In this summary, I offer this paradigm of the nature of client-centered evaluation as it is combined with practice. On the left side of this summary are short-hand expressions for what is clarified on the right side. This represents the blending of practice and evaluation processes, reflecting the central place clients should have in both. It is constructed as a kind of dialogue with the client, although I use technical terms to communicate with this professional audience:

Practice and Evaluation: General	Practice and Evaluation: Particulars
[Thorough assessment]	1." After careful consideration of the scope and nature of your (the client's) presenting concerns,
[Evidence-based general practice]	2. " and a thorough review of the relevant empirical literature that discusses how people in your situation with your kinds of concerns fared when a given intervention was used to address these concerns,
[Intervention and Evaluation Design chosen, with client-chosen goals and informed consent]	3. "I (the practitioner) have what I think are the best plans to address your concerns, as well as how I'll check with you on how well we are doing. I want to give you this information so that you can make an informed decision on whether to proceed or not. This will involve your discussing your desired outcomes regarding these concerns.
[Probable positive outcomes and possible risks]	4. "I do not know whether this intervention will work with you as it has for others with similar concerns, so I will be evaluating the positives and negatives, if any, of the outcomes of this intervention, and sharing the information with you.
[Individualized plan: Evaluation-informed specific practice]	5." I have adapted this plan of action to fit your unique circumstances, emphasizing your strengths and available social supports, and taking into consideration your problems and limitations.
[Start now, and do on-going monitoring]	6." So I propose we start now, with each of us taking parts in this process. I will carefully monitor your progress toward your desired goals. I will use this information that we collect as we go along in our contacts to fine tune the intervention. .
[Criteria for Success]	7. "We will know if these have been good choices of interventions when the concerns that brought you here are resolved, and you feel that you have attained your desired goals as much as possible. The point of this service is to give you the tools to succeed on your own."

References

Bloom, L. Z. (2008). (Im)Patient. *The seven deadly virtues and other lively essays* (pp.184-195). Columbia, SC: University of South Carolina Press.

Bloom, M., Fischer, J., & Orme, J. (2009) *Evaluating practice: Guidelines for the accountable professional* (6th ed.). Boston: Allyn & Bacon.

Hartsell, B. D. (2006). A model for ethical decision-making: The context of ethics. *Journal of Social Work Values and Ethics*, 3:1.

Harmon, S. C., Lambert, M. J., Smart, D. M., Hawkins, E., Nielsen, S. L., Slade, K., & Lutz, W. (2007). Enhancing outcomes for potential treatment failure: Therapist-client feedback and clinical support tools. *Psychotherapy Research*, 17, 379-392.

Kopels , S., & Kagle, J. D. (1993). Do social workers have a duty to warn? *Social Services Review, 67*, 101-126.

Lambert, M. J. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome

.

in routine care? *Psychological Research, 17,* 1-14.

Loewenberg, F. M., Dolgoff, R., & Harrington, D. (2000). *Ethical decisions for social work practice* (6th edition). Itasca, IL: F.E. Peacock.

Meacham, M. G. (2007). Ethics and decision making for social workers. *Journal of Social Work Values and Ethics, 4*:3.

NASW . (1999). *Code of ethics of the National Association of Social Workers*: Washington, DC: Author.

Reamer, F. G. (2006). *Social work values and ethics*. NY: Columbia University Press.

Reamer, F. G. & Shardlow, S. M. (2009). Ethical codes of practice in the US and UK: One profession, two standards. *Journal of Social Work Ethics and Values, 6*: 2.

Tarasoff v. Regents of University of California [17 Cal.3rd. 425, 1976]

Veach, R. M. (1981)*. Theory of medical ethics*. New York: Basic Books

Wilson, S. J. (1983). Confidentiality. In A. Rosenblatt & D. Waldfogel (Eds.) *Handbook of clinical social work.* San Francisco: Jossey-Bass.

Ethical Dilemmas in Sampling

Patrick Dattalo, MSW, PhD,
Virginia Commonwealth University

Abstract

This paper focuses on sampling as a nexus of ethical dilemmas experienced by social workers and other applied empirical researchers. It is argued here that social workers and other applied researchers have an ethical obligation to construct the smallest representative samples possible. Although random sampling is considered by many researchers as the gold standard methodological procedure for maximizing external validity and optimizing sample size, in practice, random sampling often is difficult to implement. Recommendations include using (1) deliberate sampling to balance a sample's composition in terms of typicalness and diversity; (2) randomization tests; (3) a broader perspective on external validity; (4) one-sided hypotheses; (5) sequential sampling; and (6) planned missing data designs.

Key Words: ethical dilemmas, external validity, unbiased sampling, optimal sample size

1.0   Introduction

The need to recognize and address ethical dilemmas at all stages of the research process is widely recognized and has been extensively discussed (cf. Engel & Schutt, 2009; English, 1997; Nelson, 1994; Rubin & Babbie, 2008). For example, ethical dilemmas may result from the researchable questions asked, the population and setting studied, the kind of information sought, and the methods used to collect data. Important ethical issues include voluntary participation and informed consent, anonymity and confidentiality, and accountability in terms of the accuracy of analysis and reporting. The need to identify and attend to the aforementioned ethical dilemmas has intensified with increasing emphasis on evidence-based practice, which is a process in which practitioners make decisions in light of the best empirical evidence available (Rubin & Babbie, 2008).

This paper focuses on sampling as a nexus of ethical dilemmas experienced by social workers and other applied empirical researchers. Emphasis will be on ethical dilemmas encountered during a quantitative research process in which the primary goal is to seek evidence about a characteristic or a relationship and to use statistical inference to generalize obtained results from a sample to a population. The National Association of Social Workers (NASW,) Code of Ethics is intended to serve as a guide to the everyday professional conduct of social workers, including researchers (NASW, 2008). At a minimum, sampling can present a conflict between social workers' ethical responsibilities for professional integrity (i.e., ensuring the rights of participants), and for evaluation and research (i.e., maximizing the generalizability of study results to the study population). The term "generalizability" is used here as a synonym for external validity, which is how well findings travel to other participants, times, and places (e.g., form a sample to a population) (Cook & Campbell, 1979).

Most social work research involves some risk or costs to participants. It is typically considered ethical if participants agree to take identified risks and to bear, or be reimbursed for, their costs. For example, risk to participants can occur if there is a new intervention, which does not address existing problems (i.e., no gain), or worse, exacerbates these problems. Costs to participants include, for example, travel and daycare expenses. Debate continues over an optimal balance between the adequate compensation of participants for their time and other expenses and their informed and voluntary decisions to be studied. Resolution of this debate seems especially difficult for participants who are vulnerable because they lack financial resources or the emotional or cognitive ability to make informed and voluntary decisions to participate in a study. Particularly vulnerable groups of study participants

include children, prisoners, persons who are terminally ill, persons who are economically disadvantaged, and persons with emotional or cognitive disabilities.

Accordingly, it is argued here that a fundamental ethical dilemma for social work researchers is to conduct studies in light of the need to maximize external validity while ensuring the rights of participants. Debate over the optimal balance between compensation and voluntary participation notwithstanding, at a minimum, social work researchers can help ensure participant rights by selecting the smallest representative sample necessary to obtain generalizable study results. For example, if a study seeks (1) sensitive information (e.g., from men in a study to compare the effectiveness of two residential substance abuse interventions); (2) information from a vulnerable population (e.g., children in a study to evaluate an intervention designed to reduce the psychosocial difficulties of children with diabetes), or (3) information during a crisis (e.g., from women seeking protective orders in cases of spousal abuse), it could be unethical to sample too many or too few people. If a sample is too small, a study could miss important effects, place unnecessary demands on participant privacy and time, or waste valuable resources. If a sample is too large, the study could make unnecessary demands on participants or misuse other resources.

Participant rights, sample size, and generalizability are interrelated issues. For clarity, the issues of ensuring participant rights and maximizing external validity and ensuring participant rights and minimizing sample size will be discussed in separate sections. Accordingly, the following discussion is focused on maximizing a sample's external validity through unbiased (i.e., random) sampling and optimal (i.e., the smallest necessary) sample size. First, the limitations of random sampling as a strategy to maximize external validity are presented. Next, three alternatives to random sampling to maximize external validity are proposed: (1) deliberate sampling for typicalness and diversity; (2) randomization tests; and (3) a broader perspective on external validity. Then, three alternatives strategies to obtain optimal sample size are presented: (1) using one-sided hypothesis; (2) sequential sampling; and (3) planned missing data designs. Please see Figure 1 for a summary of alternatives to random sampling and strategies for optimizing sample size.

Figure 1
Strategies to Maximize External Validity While Ensuring Participant Rights

Strategy	Purpose	Random Sampling Occurred?
Deliberate Sampling for Typicalness and Diversity	Reduce Sample Bias	No
Randomization Tests	Reduce Sample Bias	No
Broader Perspective on External Validity	Reduce Sample Bias	Yes and No
Use of One-Sided Hypotheses	Minimize Sample Size	Yes
Sequential Sampling	Minimize Sample Size	Yes
Planned Missing Data Designs	Minimize Time and Effort of Sample Participants	Yes

2.0   Maximizing External Validity through Unbiased Sampling

The ultimate goal of sample design is to select a set of elements from a population in such a way that descriptions of those elements accurately portray characteristics of the population (i.e., parameters) from which they were selected. Another important goal of sample design is to yield maximum precision (i.e., minimum variance) per unit cost. The sampling process begins with the identification of an appropriate population to study to answer researchable questions, and includes (1) the formulation of a sampling strategy, and (2) determination of sample size and composition to maximize the external validity.

2.1   What Random Sampling Does and Does Not Do

The term random sample, also called probability sample, is used to describe a sample that is free from systematic error. A sample is unbiased, then, if every element in a population has an equal chance of being selected. According to classical statistical sampling theory, if random selection from a known population is performed, characteristics of the sample can be inferred and tend to mirror corresponding characteristics of the population. If random sampling is not performed, there is no theoretical basis for statistical inference. Only information about a sample can be described. However, although random sampling for representativeness minimizes systematic error, sampling biases still can occur for the following reasons:

1. A complete randomization process is usually not implemented (Cook, 1993). Frequently, only units are randomized, which is only one of three different areas that define an event. The other two areas that define

an event are place or setting, and time (Cook & Campbell, 1979);

2. Random sampling does not minimize all error in a research design. There are other types of bias in the sample that may contribute to error, such as non-sampling bias (e.g., measurement error (Henry, 1990);

3. The sample may not be representative of the population because it is too small, and therefore, likely to be too homogeneous;

4. The representativeness of the sample may be impacted by attrition and refusal of the participants to take part in a study;

5. Random sampling permits causal generalization to a target population but not *across* multiple populations (Cook & Campbell, 1979). The latter is important for establishing an abstract principle of causality and is best done through multiple replications across units, setting, and time;

6. Conclusions drawn from random samples are based on inferential theory or the probability of the occurrence of an event. Random sampling, alone, does not guarantee accurate estimates of population parameters; and

7. Random sampling usually requires considerable resources compared with nonrandom sampling strategies. If resources are inadequate to enumerate a representative sampling frame and to draw sample size for adequate statistical power, the hoped for goal of random sampling (i.e., an unbiased sample), may not be achieved.

If random sampling is not possible social work researchers should consider a nonprobability alternative, such as deliberate sampling, to help to ensure generalizabiltiy.

2.2 Deliberate Sampling for Typicalness and Diversity

Cook and Campbell (1979) argue that deliberate sampling, also termed purposive sampling, may be useful if a sample is carefully constructed. Deliberate sampling is a type of nonprobability sampling in which elements are knowingly chosen based on a study's research questions (Cook & Campbell, 1979). Blankertz (1998) and Bull (2005) emphasize the following two variations on deliberate sampling: (1) deliberate sampling for diversity, which involves selecting a sample with a wide range of characteristics that are expected to influence results; and (2) deliberate sampling for typical instances, which involves selecting at least one instance of each class that is impressionistically similar to that class's mode.

Deliberate sampling can be used to achieve a variety of research goals. Potential applications of deliberate sampling are as follows:

1. Study a time-limited population (e.g., all clients being served by a department of social services);

2. Study a subset of a population (e.g., only clients being provided child protective services by a department of social services);

3. Primary data analysis; that is, deliberate sampling can be used to select clients for a pilot study that will be used to guide a larger scale study;

4. Secondary data analysis; that is, to select a sample from an existing data set for a secondary analysis; and

5. Descriptive analysis; that is, a researcher can select a small subsample and closely examine typical and unusual or extreme elements.

Deliberate sampling shares certain characteristics with stratified sampling. In a stratified sample, the sampling frame is divided into non-overlapping groups or strata (e.g., age groups, gender). Then, a random sample is taken from each stratum. Stratified sampling uses groups to achieve representativeness, or to ensure that a certain number of elements from each group are selected. Like stratified random sampling, deliberate sampling can be used to control the characteristics of cases being selected (cf. Armitage, 1947; Kott, 1986).

There is empirical evidence of the ability of stratified random sampling to increase precision when the strata have been chosen so that members of the same stratum are as similar as possible in respect of the characteristic of interest; the larger the differences between strata, the greater the gain in precision (cf. Armitage, 1947; Kott, 1986). Stratification (and deliberate sampling) can help to ensure that not only the overall population, but also that key subgroups of the population, are represented. For example, if the subgroup is small, and different sampling fractions are used to "over-sample the small

group" stratified random sampling will generally have more statistical precision than simple random sampling. The benefits of stratification are greatest when the strata or groups are homogeneous; that is when within-groups variability is lower than the variability for the population.

The following example seeks to achieve the first of the five aforementioned research goals (i.e., study a time-limited population) and follows a procedure described in Blankertz (1998). Note that this procedure is versatile and could be used to achieve any of these five research goals.

A researcher conducts a study to determine the effectiveness of a peer-led eating disorders prevention intervention in reducing eating-disorder risk factors in young women (18 to 21 years of age). The intervention is implemented, at the discretion of the school, in public four-year colleges and universities in a state as a part of new student orientation. First, a sampling frame is used to conceptualize a deliberate sample for diversity and typicalness of new students. Next, all new students in this deliberately constructed sampling frame are randomly assigned to either an intervention or a control group. The intervention consists of eight two-hour group sessions that were delivered by trained peer facilitators. Participants completed questionnaires that assessed eating-disorder risk factors pre and post treatment with higher scores indicating a greater risk of eating disorders. Demographic characteristics of participants, including age, gender, race/ethnicity, region (based on the home address of students), and BMI (Body Mass Index) score were also collected. Results consisted of a comparison of the intervention and control group means. Further analysis consisted of a comparison of the intervention and the control groups in three subsamples (i.e., two for diversity and one for typicalness).

Deliberate sampling for diversity involves selecting two subsamples, each chosen to "differ as widely as possible from each other" (Cook & Campbell, 1979, p. 78). In addition, these two subsamples samples should be selected to vary across several characteristics, including time and place. It is helpful to view each characteristic in each subsample as the endpoint on a continuum of a ratio, interval, or ordinal variable. For nominal variables, each characteristic represents a different category of that variable. Each subsample should contain clusters of elements that represent endpoints of ordinal, interval, or ratio variables, or the different categories of a nominal variable. That is, each subsample should contain values that represent opposing endpoints or categories.

Cook and Campbell (1979) explain that "given the negative relationship between 'inferential power' and feasibility, the model of heterogeneous instances (i.e., sampling for diversity)would seem most useful, particularly if great care is made to include impressionistically modal (i.e., typical) instances among the heterogeneous ones" (p. 78). Moreover, Cook and Campbell (1979) conclude that

Practicing scientists routinely make causal generalizations in their research, and they almost never use formal probability sampling when they do. Scientists make causal generalizations in their work by using five closely related principles: (1) surface similarity, (2) ruling out irrelevancies, (3) making discriminations, (4) interpolation and extrapolation, and (5) causal explanation. Deliberate or purposive sampling for heterogeneous instances; and impressionistic or purposive sampling of typical instances are essential components of these principles (p. 24).

With the careful matching of sampling strategy to purpose, deliberate sampling can be a useful alternative to random sampling. If random sampling is not possible social work researchers should consider a nonprobability alternative, such as randomization tests, to help to ensure generalizabiltiy.

2.3   Randomization Tests

According to Howell (2007), randomization tests differ from parametric tests as follows:

1.   There is no requirement that a sample is randomly drawn from a population;
2.   There is no assumption about the population from which the sample is drawn (e.g., it is normally distributed), although as sample size increases, the distribution produced by permutations approaches the normal distribution;
3.   Because there are no assumptions about a population, no sample statistics are used to estimate population parameters; and
4.   Although test statistics are calculated, they are not utilized in the same way as they are in parametric hypothesis testing. Instead, the data are repeatedly

randomized across groups, and test statistics are calculated for each randomization. Therefore, at least as much as parametric tests, randomization tests emphasize the importance of random assignment of participants to treatments.

A randomization test can be described as follows. A test statistic is computed for study data (e.g., a *t*-test), termed an obtained result. Then, these data are permuted repeatedly and the test statistic is computed for each of the resulting data permutations. When data are permuted, the sample is divided or rearranged by random assignment without replacement to fill the first group, and then to fill the second group until each group contains a new sample of the same size as the original group. These permutations, including the one representing the obtained result, constitute the reference set for determining significance. The proportion of data permutations in the reference set that have a test statistic values greater than, or for certain test statistics, less than or equal to the value for the obtained result, is the p-value.

For example (hypothetical), in a study of the effectiveness of a new treatment to increase empathy in a group of spouse abusers, participants are randomly assigned to either a treatment or a control group. One group is a control condition with scores of 25, 22,23,21,17 on an empathy scale, and the other group was the treatment condition with scores of 30, 27, 28, 29, 29. If the treatment had no effect on scores, the first number that was sampled (25) would be equally likely to occur in either group. With five observations in each group, and if the null hypothesis is true, any five of these 10 observations would be equally likely to occur in either group. These data are "exchangeable" between conditions.

After calculating all of the possible arrangements of the aforementioned 10 observations with five observations in each group (there are 252 possible arrangements), the relevant test statistic is calculated (independent groups *t*-test) for each possible arrangements, and compared to the obtained *t*-test value (4.9252) to test the null hypothesis of no difference in scores between the treatment and the control group. In this case, there are two arrangements of these data that would have a smaller mean for the control group and a larger mean for treatment group. For a one-tailed test, there are two data sets that are at least as extreme as these data. Consequently, a difference that is at least as large as the obtained *t*-value of 4.9252 would occur two times out of 252 for a probability of .006 under the null hypothesis. That is, this difference is statistically significant at $p < .01$. Edgington (2007), Erceg-Hurn and Mirosevich (2008), and Rodgers (1999) provide more detailed explanations and examples of the use of randomization tests.

Stata (http://www.stata.com/) is a commercial general purpose statistical software that can be use to perform a wide range of randomization tests. A free alternative is David Howell's program, Resampling.exe, which is available online from http://www.uvm.edu/~dhowell/StatPages/Resampling/Resampling.html. This software can be used to perform a limited range of randomization tests.

2.4 Toward a Broader Perspective on External Validity

Whether or not random sampling is possible social work researchers should consider a broad perspective on external validity. Reasoning from data points in a sample to an estimate of a population characteristic is an instance of induction. Hume, who was an 18th century Scottish philosopher, usually is credited with discovering "the problem of induction." As identified by Hume, the problem of induction is how to establish induction itself as a valid method for empirical inquiry. See, for example, Wood (2000) for a detailed explanation. According to Rosenberg (1993),

> Hume recognized that inductive conclusions could only be derived deductively from premises (such as the uniformity of nature) that themselves required inductive warrant, or from arguments that were inductive in the first place. The deductive are no more convincing than their most controversial premises and so generate a regress, while the inductive ones beg the question. Accordingly, claims that transcend available data, in particular predictions and general laws, remain unwarranted (p. 75).

To clarify the fundamental limitations of statistical, sampling-based generalization, consider the hypothesis, H_A that the average difference in the perceived effectiveness, which two groups of social workers (i.e., those working in a public social services agency and those working in a public mental health agency) associate with a particular intervention, is 3. In other words, a researcher does not know the numerical value of the average difference in perceived effectiveness between two groups, but hypothesizes it to be 3 (where effectiveness is measured, for example, on a scale from 1 to 5). The researcher then tests the H_0 hypothesis that there is no difference between the

average perceived effectiveness between the two groups by taking a random sample of social workers from each group, and uses the average of the sample from each group as an estimate of the average of the perceived effectiveness for that group.

If the legitimacy of inductive reasoning is unquestioned, then the researcher could reason that the sample average is generalizable to the population average. However, if the legitimacy of inductive reasoning is questioned and Hume's argument is applied, there would be no sound basis for making any statement about the value of the population average. This idea can be expressed as follows: Just because all differences between the two groups in past samples have an average of 3 does not mean that all or any differences between the two groups in future samples will have an average of 3.

Significance tests based on probability sampling, at best, provide very specific information about a population based on a sample's characteristics. In statistical significance testing, the p-value is the long-run probability of obtaining a result (e.g., differences in perceived effectiveness between two groups) at least as extreme as the given result, assuming the null hypothesis. As Cohen (1994) pointed out, what researchers and consumers of research want to know is the population parameter, given the statistic in the sample and the sample size. Unfortunately, the direction of the inference is from the population to the sample, and not from the sample to the population (Thompson, 1997). That is, the logic of hypothesis testing assumes the null is true in a population, and asks: given this assumption about the parameters of a population, what is the probability of the sample statistic?

Campbell and Stanley (1966) eloquently call attention to the "painful" limitations of inductive reasoning when they state:

> Whereas the problems of internal validity are solvable within the limits of the logic of probability statistics, the problems of external validity are not logically solvable in any neat, conclusive way. Generalization always turns out to involve extrapolation into a realm not represented in one's sample. Such extrapolation is made by assuming one knows the relevant laws. Thus, if one has an internally valid [design], one has demonstrated the effect only for those specific conditions which the experimental and control group have in common, i.e., only for pretested groups of a specific age, intelligence, socioeconomic status, geographical region. . . Logically, we cannot generalize beyond these limits; i.e., we

cannot generalize at all. But we do attempt generalization by guessing at laws and checking out some of these generalizations in other equally specific but different conditions. In the course of the history of a science we learn about the "justification" of generalizing by the cumulation of our experience in generalizing, but this is not a logical generalization deducible from the details of the original experiment. Faced by this, we do, in generalizing, make guesses as to yet unproven laws, including some not even explored (p. 17).

Campbell and Stanley (1966) conclude that "induction or generalization is never fully justified logically" (p. 17), and they argue that a sample can, at best, offer only limited support for generalization.

Evidence of result generalizability is critical to the accumulation of knowledge, and should be provided by authors. Accordingly, social work researchers should always provide a detailed description of a study's sample. A detailed description is necessary to understand the population being studied and to judge whether the extent of generalizing results seems appropriate. Also, when possible, a comparison of study participants and information about the population should be provided to enable readers to evaluate a sample's representativeness in terms of the larger population from which it was drawn.

The ability to generalize from one situation to another depends on the ability to understand underlying principles and to recognize which underlying principles apply in a given situation. According to Mook (1983), there is no alternative to thinking through, case by case (1) what conclusions are desired; and (2) whether the specifics of a sample or setting prevent these conclusions (p. 386). Mook argues that any generalization to a population of interest must be made on other than statistical grounds.

A broader perspective on generalization recognizes that it requires a series of inferences and judgments regarding the appropriateness of applying findings, concepts, or theories to new or different settings or phenomena. Generalization, therefore, involves identifying similarities and differences between research participants and between research contexts to assess whether a finding or theory is relevant to a new domain (Audi, 2003). Lee and Baskerville (2003) propose a framework of four different types of generalizability built upon the distinction between empirical and theoretical statements as the inferential content. Empirical statements refer to data from and descriptions of

empirical phenomena; theoretical statements refer to phenomena that cannot be directly observed and therefore can only be theorized from empirical data or other theories (p. 232). A second distinction that forms this typology is the distinction between "generalizing *from*" and "generalizing *to*" (p. 232).

Generalization is usually considered to be the ultimate goal of quantitative research. However, an expanding acceptance of the complementary and supplementary roles of qualitative and quantitative approaches to social work should serve as a reminder of the need to recognize the tension between the particular and the general throughout the research process, and of the potential contributions of both random and nonrandom sampling strategies. This tension suggests the importance of thinking more deeply about the content, function, and ethical implications of result generalizations.

3.0 Maximizing External Validity through Optimal Sample Size

Sample size influences the quality and accuracy of empirical research. In general, increased sample size is associated with decreased sampling error. The larger the sample, the more likely the results are to represent the population. However, the relationship between sampling error and sample size is not simple or proportional. There are diminishing returns associated with adding elements to a sample. The relationship between sample and accuracy may be clarified by the use of the concept of statistical power. The notion of statistical power is attributed to Neyman and Pearson (1928), although Fisher (1925) addressed similar issues in his discussions of test and design sensitivity and popularized in the behavioral sciences by Jacob Cohen (c.f., Huberty, 2002, for a detailed discussion).

Power is the probability of rejecting the null when a particular alternative hypothesis is true. More simply, statistical power is the probability of detecting a pre-specified effect size (e.g., a minimally important one). An underpowered study is one for which the projected scientific or clinical value is unacceptably low because it has less than 80% chance of resulting in statistical significance at an a priori set a level (usual $p < .05$). Researchers should avoid conducting studies that are ''underpowered.'' Conversely, researchers should avoid conducting studies with too large a sample size. Studies with samples that are too large may needlessly place respondents at risk, waste their time, and misuse other resources, such as professional time and scarce research dollars. Accordingly, researchers should focus on determining the smallest necessary sample size.

Bacchetti, Wolf, Segal, and McCulloch (2005a; 2005b) discuss how sample size influences the balance that determines the ethical acceptability of a study. That is, the balance between the burdens that participants accept and the clinical or scientific value that a study can be expected to produce. The average projected burden per participant remains constant as the sample size increases, but the projected study value does not increase as rapidly as the sample size if it is assumed to be proportional to power or inversely proportional to confidence interval width. This implies that the value per participant declines as the sample size increases and that smaller studies therefore have more favorable ratios of projected value to participant burden. Bacchetti et al. (2005a; 2005b) provocatively conclude that their argument ''does not imply that large studies are never ethical or that small studies are better, only that a small study is ethically acceptable whenever a larger one is'' (p. 113).

Analysis by Bacchetti et al. (2005a; 2005b) addresses only ethical acceptability, not optimality; large studies may be desirable for other than ethical reasons. The balance point between burden and value cannot be precisely calculated in most situations because both the projected participant burden and the study's projected value are difficult to quantify, particularly on comparable scales. Bacchetti et al. (2005a; 2005b) provided a service by encouraging researchers to think of value and burden on a per-participant basis and by arguing that the expected net burden per participant may often be independent of sample size (Prentice, 2005). Institutional review boards are becoming more sophisticated regarding power and sample size issues, and, consequently, there could be fewer studies with inappropriate (too large or too small) sample sizes in the future. If random sampling is possible, social work researchers should consider testing a one-sided hypothesis to minimize sample size.

3.1 One- Versus Two-Sided Hypotheses as Determinants of Sample Size

The estimation of the minimum sample size requires the specification of the minimal difference in outcome (effect size or δ) that would be practically important to be detected. In addition, researchers must specify (1) an acceptable α-level ($1 - \alpha$ is the probability of detecting a significant difference when the treatments are really equally effective), (2) an acceptable β-level ($1 - \beta$ is the probability of not detecting a significant difference when there really is a difference of magnitude δ or larger), and (3) the

standard deviation of the hypothesized effect size in the population of interest. Finally, a researcher should explicitly choose between two-sided or one-sided statistical testing. As Knottnerus and Bouter (2001) suggest, the importance of this last criteria decision often is neglected. A one-tailed hypothesis specifies a directional relationship between groups. That, is the researcher not only states that there will be differences between the groups but specifies in which direction the differences will exist. Anytime a relationship is expected to be directional (i.e., to go one specific way) a one-tailed hypothesis is being used. This is the opposite of a two-tailed hypothesis. With a two tailed hypothesis the researcher would predict that there was a difference between groups, but would make no reference to the direction of the effect (Bland & Altman, 1994).

Knottnerus and Bouter (2001) argue that a research hypothesis expresses scientific uncertainty regarding a plausible, potentially practically important effect. Consequently, a research question is often hypothesis-driven and typically "one-sided." Accordingly, if a new intervention is compared with no treatment, the one-sided approach would be adequate. Moreover, for example, assuming $\alpha = 0.05$, $\beta = 0.80$, a moderate effect size of Cohen's d, and equal numbers in the intervention and treatment groups of a study, each group needs a minimum sample size of 88 in case of one-sided testing; and 105 per group in case of two-sided testing. This means that the two-sided approach requires an additional 34 or 19% more participants than the one-sided approach.

As Moye and Tita (2002) explain, however, there are important limitations to the one-tailed test in a clinical research effort. A major difficulty is that the one-sided testing philosophy reveals a potentially dangerous level of investigator consensus that there is no possibility of participant harm produced by the intervention being tested. Although the two-sided hypothesis test can complicate experimental design, increasing sample size requirements, this approach is ultimately more informative and potentially prevents subsequent exposure of research participants and the general population to harmful interventions. Although two-sided tests are only capable of establishing a difference (rather than a difference and direction), researchers may explore their data and determine in which direction any significant difference lies. In fact, researchers should routinely report a confidence interval around an outcome measure, such as a mean.

When deciding whether a one- or two-sided hypothesis approach is most appropriate for a study that they are planning, researchers may consider prior evidence and the practice implications of the intervention being studied (Enkin, 1994). This body of existing studies, including meta-analyses may provide evidence in support of these methodological choices. For interventions not previously studied or about which few studies have been conducted, a one-sided view seems reasonable if the comparison is between that intervention and no intervention. Regardless of which hypothesis testing approach is selected, the researcher should formulate the research hypothesis *a priori*. If random sampling is possible, another strategy to minimize sample size is sequential sampling (Dunnett & Gent, 1996; Posch & Bauer, 2000; Whitehead, 1997).

3.2   Sequential Sampling

Sampling strategies, whether probability or non-probability, can be categorized as either single, (also termed fixed), or multiple, (also termed sequential) (Stephens, 2001). With a sequential sampling strategy, after a first sample is tested, there are three possibilities: accept, reject, or make no decision about a hypothesis. If no decision is made, additional samples are collected and each sample is analyzed to determine whether to accept or reject a hypothesis, or to proceed and collect another sample. More specifically, in a sequential sampling design, data are analyzed periodically, and sample size is not a single fixed number. An appropriate schedule for interim analyses is established together with a stopping rule, which defines the outcomes that lead to early termination of a study.

The classical theory of hypothesis testing is based on a sample of fixed size (Neyman & Pearson, 1928). In this sample, the null hypothesis H_0 is tested against an alternative hypothesis H_1. A significance level α is defined a priori (i.e., in advance of data collection), which is the probability of the null hypothesis being falsely rejected. Consequently, in a classical fixed sample design, the sample size is set in advance of data collection, and hypothesis testing occurs after all observations have been made. The main design focus is on choosing a sample size that allows a study to discriminate between H_0 and H_1 and answer the research questions of interest.

In fixed sample design, then, together with practical considerations, a study's sample size is determined a priori by setting up null and alternate hypotheses concerning a primary parameter of interest (θ), and then specifying a Type I error rate (α) and power (1-B) to be controlled at a given treatment effect size ($\theta = \Delta$). Usually, traditional values of α and B are used (i.e.., $\alpha = .05$, $B = .20$); however, there can be considerable debate over the choice of the effect size (Δ). In general, the smaller the effect size, the larger the sample size needed to

detect it. The choice of Δ is crucial because, for example, reducing a selected effect size by 50% leads to a quadrupling in the sample size for a fixed sample. Using a sample size that is small relative to a selected effect size can result in a study that is underpowered (i.e., unlikely to detect a smaller, but possibly still important, effect). Consequently, Cohen (1988) and others (cf. Adcock, 1997; Orme & Hudson, 1995; Stolzenberg & Relles, 1997) have proposed the use of a sample big enough to detect the smallest worthwhile effect. A disadvantage of all fixed sample designs is that estimated sample size is the same regardless of the magnitude of the true intervention effect. Accordingly, one approach to increasing the congruence between estimated and true effect sizes is to perform interim analyses with sequential sampling.

With a sequential sampling strategy, after a first sample is tested, there are three possibilities: accept, reject, or make no decision about a hypothesis. If no decision is made, additional samples are collected and each sample is analyzed to determine whether to accept or reject a hypothesis or to proceed and collect another sample (Jennison & Turnbull, 2000). More specifically, in a sequential sampling design, data are analyzed periodically, and sample size is not a single fixed number. An appropriate schedule for interim analyses is defined together with a stopping rule, which defines the outcomes that lead to early termination of the study. For example, sequential sampling allows consecutive testing, with possible rejection of the null hypothesis, after each set of observations in a pair of groups (e.g., intervention and control).

With sequential sampling, for ethical and practical reasons, results can be monitored periodically and, if sufficiently large or small effects are observed, data collect may be stopped early. Evidence suggests that sequential designs require fewer participants than fixed sampling designs (Jennison & Turnbull, 2000; Whitehead, 1997). Tests of sequential samples have been developed that allow for early stopping to reject or accept the null hypothesis while preserving the integrity of the test; that is, maintain desired Type I error and power.

Sequential sampling design parameters include (1) power; (2) sample size; (3) number and timing of analyses; (4) criteria for early stopping (i.e., evidence against the null hypothesis, the alternative hypothesis, or both); and (5) stopping rules (i.e., the relative ease or conservatism with which a study will be terminated at the earliest analysis versus later analyses). A sequential sampling plan consists of two or more stopping rules. Data are monitored at interim time-points and the process is terminated early if, for example, a difference between two interventions in

terms of an outcome can be established statistically at any one of the interim looks. Since the data will be tested repeatedly in a group-sequential study, the burden of proof must be more stringent at each of the interim looks than without interim monitoring. Otherwise, there is an increased that chance fluctuations in the data will be misinterpreted as demonstrating a real underlying effect. This increasing stringency is accomplished by establishing a stopping boundary at each interim look (Pampallona & Tsiatis, 1994; Proshan & Hunsberger, 1995).

In summary, sample size estimation is a key component of empirical research. A sequential sampling strategy may be most useful when appropriate effect sizes and estimates of variability necessary for sample size calculations are not known. In addition to saving time and resources, sequential sampling can reduce study participants' exposure to an inferior intervention. Sequential sampling also may be useful when conducting a pilot study. Sequential sampling helps determine whether the researcher has taken a large enough pilot sample to properly evaluate different sampling designs, and to use the standard deviation from the pilot sample to calculate sample size for a larger scale study.

A limitation of group-sequential sampling is an increased probability of Type I error because of repeated significance testing. Unadjusted, repeated significance testing of the accumulating data increases the overall significance level beyond the pre-specified nominal significance level. Consequently, interim analyses and their interpretations need to be done judiciously. To reduce the probability of Type I error, a study's protocol should contain a formal rule for stopping the study early. The decision to conduct an interim analysis should be based on sound scientific reasoning. Researchers should avoid the use of vaguely defined and misunderstood terms and phrases such as "administrative looks," "administrative interim analyses," "interim analysis for safety, and "interim analysis for sample size adjustment" (Sankoh, 1999). If random sampling is possible, a third strategy that social work researchers should consider using is planned missing data designs to minimize measurement instrument length, and consequently, costs to both participants and investigators (Graham et al., 2006).

3.3   Planned Missing Data Designs

When researchers design measurement instruments for a study, they universally must balance a desire to seek information from participants against participants' costs of providing this

information. Graham et al. (2006) described a planned missingness design called two-method measurement (also see Allison & Hauser, 1991, who describe a related design). For example, social work researchers typically (1) collect demographic and other background information data; and (2) and administer at least one standardized scale of moderate length. The two-method measurement design may allow the researcher to collect complete demographic data, and partial data (on a random sample of participants) for the standardized scale(s). A possible limitation of this type of design is that it requires the use of structural equation modeling (Muthe´n, Kaplan, & Hollis, 1987).

Another, design described by Graham et al. (2006) for maximizing information while minimizing participant costs is the three-form design. In its generic form, the three-form design allows researchers to increase by 33% the number of questions for which data are collected without changing the number of questions asked of each participant by dividing all questions asked into four items sets. One set (X) contains questions most central to the study outcomes, and is asked of all participants. Three additional sets of questions (A, B, C) are constructed, with each set containing one-third of the remaining question. Sets A, B, and C are rotated, such that one set is omitted from each of the three forms (i.e., X and two of the A, B, C sets).

An advantage of planned missing data designs is that less data are required, and therefore, less data needs to be collected. However, it is not clear how patterns of incomplete data should be structured and incorporated into research designs, particularly in longitudinal designs. In cross-sectional research designs, the use of a reference variable has been shown to be effective in terms of obtaining the correct estimates in the context of planned incomplete data structures (McArdle, 1994; Graham, Hofer, & MacKinnon, 1996). In such designs, a reference variable refers to obtaining complete data for a given variable, or for one variable within each factor of the research design. A reference variable is used as a baseline measure, which should aid the imputation process since full information is provided for all participants in relation to other variables being studied. The use of a reference variable also would seem to be an attractive option in some longitudinal research designs, since it could be incorporated across administrations. The efficacy of this design was tested by Bunting and Adamson (2000) through a series of simulations, and the result suggested that parameter estimates are both precise and efficient.

4.0 Conclusions

The need to recognize and address ethical dilemmas at all stages of the social work research process is widely recognized. This paper has focused on sampling as a nexus of ethical dilemmas. It has been argued that social workers and other applied researchers have an ethical obligation to construct the smallest representative samples possible. Random sampling is considered by many researchers as the gold standard methodological procedure for maximizing external validity and optimizing sample size. However, in practice, random sampling often is difficult to implement. Although assembling the smallest representative sample possible may seem daunting at times, recommendations include (1) deliberate sampling to balance a sample's composition in terms of typicalness and diversity; (2) randomization tests; (3) a broader perspective on external validity; (4) use of one-sided hypotheses; (5) sequential sampling; and (6) planned missing data designs. Moreover, fulfilling the aforementioned ethical obligation to construct the smallest representative samples possible usually will benefit from a mix of strategies to maximize external validity and minimize sample size.

References

Adcock, C. J. (1997). Sample size determination: A review. *Statistician*, *46* (2), 261–283.

Allison P. D., & Hauser, R. M. (1991). Reducing bias in estimates of linear models by remeasurement of a random subsample. *Sociological Methods & Research. 19*, 466–492.

Armitage, P. (1947). A comparison of stratified with unrestricted random sampling from a finite population. *Biometrika*, *34*, 273-280.

Audi, R. (2003). *Epistemology: A contemporary introduction to the theory of knowledge.* New York: Rutledge.

Bacchetti, P., Wolf, L. E., Segal, M. R., & McCulloch, C. E. (2005a). Ethics and sample size. *American Journal of Epidemiology*, *161*, 105–110.

Bacchetti, P., Wolf, L. E., Segal, M. R., & McCulloch, C. E. (2005b). Bacchetti, et al. respond to ''Ethics and sample size—Another view.'' *American Journal of Epidemiology*, *16*, 113.

Bland, J. M., & Altman, D. G. (1994). One and two sided tests of significance. *British Medical Journal, 309*, 248.

Blankertz, L. (1998). The value and practicality of deliberate sampling for heterogeneity. *American Journal of Evaluation*, *19*, 307-324.

Bull, B. (2005). Exemplar sampling: Nonrandom methods of selecting a sample which characterizes a finite multivariate population. *The American Statistician*, *59*, 166-172.

Bunting, B. P., & Adamson, G. (2000). Assessing reliability and validity in the context of planned incomplete data structures for multitrait-multimethod models. *Developments in Survey Methodolgy.* Available online http://mrvar.fdv.unilj.si/pub/mz/mz15/bunting.pdf

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Cohen, J. (1994). The earth is round (p < .05). *American Psychologist*, *49*, 997-1003.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Cook, T. D. (1993). A quasi-sampling theory of the generalization of causal relationships. In L. B. Sechrest & A. G. Scott (Eds.), *Understanding causes and generalizing about them* (pp. 39-83). San Francisco: Jossey-Bass.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation – Design & analysis issues for field settings*. Boston: Houghton Mifflin Company.

Dunnett C., & Gent, M. (1996). An alternative to the use of two-sided tests in clinical trials. *Statistics in Medicine*, *15*, 1729–1738.

Edgington, E. S. (2007). *Randomization tests*. New York: Marcel Dekker.

Engel, R., & Schutt, R. K. (2009). *The Practice of Research in Social Work.* Thousand Oaks, CA: Sage.

English, B. (1997). Conducting ethical evaluations with disadvantaged and minority target groups. *Evaluation Practice*, *18*, 49-54.

Enkin, M. W. (1994). One and two-sided tests of significance: One-sided tests should be used more often. *British Medical Journal*, *309*, 873-874.

Erceg-Hurn, D.M., & Mirosevich, V.M. (2008). Modern robust statistical methods: An easy way to maximize the accuracy and power of your research. *American Psychologist*, *63,* 591-601.

Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh: Oliver and Boyd.

Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods*, *11*, 323–343.

Graham, J. W., Hofer, S. M., & MacKinnon, D. P. (1996). Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research*, *31*, 197-218.

Henry, G. T. (1990). *Practical sampling*. Newbury Park, CA: Sage.

Howell, D. (2007). *Randomization tests on correlation coefficients*. Available online at http://www.uvm.edu/~dhowell/StatPages/Resampling/RandomCorr/randomization_Correlation.html**.**

Huberty, C. J. (2002). A history of effect size indices. *Educational and Psychological Measurement*, *6*, 227-240.

Jennison, C., & Turnbull, B. W. (2000). *Group sequential methods with applications to clinical trials*. Boca Raton, FL: Chapman & Hall.

Knottnerus, J.A., & Bouter L.M. (2001). The ethics of sample size: Two-sided testing and one-sided thinking. *Journal of Clinical Epidemiology*, *54*, 109-110.

Kott, P. S. (1986). Some asymptotic results for the systematic and stratified sampling of a finite population. *Biometrika*, *73*, 485-491.

Lee, A. S., & Baskerville, R. L. (2003). Generalizing generalizability in information systems research. *Information Systems Research*, *14*, 221–243.

McArdle, J. J. (1994). Structural factor analysis experiments with incomplete data. *Mutivariate Behavioral Research, 29*, 409–54.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, *38*, 379–387.

Moye, L. A., & Tita, A. T. N. (2002). Defending the rationale for the two-tailed test in clinical research. *Circulation*, *105*, 3062-3065.

Muth´en, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika, 52,* 431–462.

NASW. (2008). The NASW code of ethics. Available online at http://www.socialworkers.org/pubs/code/code.asp

Nelson, J. C. (1994). Ethics, gender, and ethnicity in single-case research and evaluation. *Journal of Social Service Research*, *18*, 139-151.

Neyman, J., & Pearson, E. S. (1928). On the use and interpretation of certain test criteria for

purposes of statistical inference. Part I and II. Biometrika, *20,* 174-240, 263-294.

Orme, J. G., & Hudson, W. W. (1995). The problem of sample size estimation: Confidence intervals. *Social Work Research*, *19*, 121–127.

Pampallona, S., & Tsiatis A. A. (1994). Group sequential designs for one-sided and two-sided hypothesis testing with provision for early stopping in favor of the null hypothesis. *Journal of Statistical Planning and Inference*, *42*, 19-35.

Posch, M., & Bauer, P. (2000). Interim analysis and sample size reassessment. *Biometrics*, *56*, 1170–1176.

Prentice, R. (2005). Invited commentary: Ethics and sample size—Another view. American *Journal of Epidemiology*, *161*, 111–112.

Proschan, M. A., & Hunsberger, S. A. (1995). Designed extension of studies based on conditional power. *Biometrics*, *51*, 1315–1324.

Rodgers, J. L. (1999). The bootstrap, the jackknife, and the randomization test: A sampling taxonomy. *Multivariate Behavioral Research*, *34*, 441-456.

Rosenberg, A. (1993). Hume and the philosophy of science. In D. Norton (Ed.), *The Cambridge companion to Hume* (pp. 64-89). New York: Cambridge University Press.

Rubin, A., & Babbie, E. R. (2008). *Research methods for social work*. Belmont, CA: Brooks/Cole.

Sankoh, A. J. (1999). Interim analyses: An update of an FDA reviewer's experience and perspective. *Drug Information Journal, 33*, 165-176.

Stephens, K. M. (2001). *The handbook of applied acceptance sampling*. Milwaukee, Wisconsin: ASQ Quality Press.

Stolzenberg, R. M., & Relles, D. A. (1997). Tools for intuition about sample selection bias and its correction. *American Sociological Review*, *62*, 494–507.

Thompson, B. (1997). Statistical significance testing practices in *Journal of Experimental Education*, *66*, 75-83.

Whitehead, J. (1997). *The design and analysis of sequential clinical trials*. New York: John Wiley.

Wood, A. (2000). Hume: The problem of induction. Available online at http://www.stanford.edu/~allenw/Phil102/Hume%20-%20Induction.doc

Forum Article

Teaching Social Work Values and Ethics Online

Stephen M. Marson, Ph.D. University of North Carolina at Pembroke
Guo Wei, Ph.D. University of North Carolina at Pembroke
Barbara M. Marson, Ph.D. East Carolina University

Abstract

During the academic year 2004-2005, SWK 4500 Social Work Values and Ethics was taught both in a face-to-face format and on Blackboard. In the Blackboard class, n = 23; while in the face-to-face class, n = 16. The professor, content and syllabi for both courses were identical. Most importantly, **both** classes took the identical final exam on **Blackboard**. An unmatched t-test indicated that there was a significant difference between exam scores [t = 2.42; df =18; p < .026]. An item analysis of the exam uncovered that the Blackboard students were troubled with the test items especially addressing the application of abstract ethical concepts to practice situations. Separate factor analyses for the two groups found that the performance of face-to-face students differed primarily on the understanding of ethical theory while Blackboard students were widely divided by their knowledge for ethical theory as well as applying ethics to practice situations. Thus, the factor analysis supported the t-test results. The final was the only test given and there were no differences for any other class assignments.

Key Words: Blackboard, face-to-face, final exam, unmatched t-test, factor analysis, social work

1. INTRODUCTION

1.1. Blackboard

Blackboard is a registered and copyrighted internet protocol and software that provides a framework for teaching a course online [see: http://www.blackboard.com/us/index.aspx]. Professors may employ Blackboard in two ways. First, professors can utilize Blackboard to enhance classroom (face-to-face) teaching. Professors can email assignments, notes, display figures, assess who missed class, give exams, etc. When the professor is unavailable because of illness or a conference, class lectures and class discussions including digital films may be offered on Blackboard participation as an acceptable substitute for a classroom appearance. In this manner, Blackboard enhances classroom instruction.

The second manner in which Blackboard is employed includes conducting an entire course without a physical classroom or seeing the students face-to-face. The Blackboard stakeholders assert that everything that can be presented in a face-to-face classroom can be presented via Blackboard. Students can find the syllabus, tests, communication centers, discussion boards, videos, assignments, etc. on Blackboard. Professors can present the class in a synchronous or asynchronous manner. Each succeeding version of Blackboard has offered greater technical sophistication than previous versions. Within this study, version 6.3.1.424 was employed.

1.2. The Study

This research addresses the learning outcomes of a Blackboard course and a face-to-face course. The research question is: "Is there a significant difference between identical final exam scores for Blackboard course and face-to-face students?" Thus, the t-test null hypothesis would be: "There is no difference between final exam scores for Blackboard course and face-to-face students." Using statistical symbols, the null hypothesis would be stated as:

$$H_0 : \mu_1 = \mu_2$$

where μ_1 and μ_2 represent the average final exam scores for the Blackboard class and face-to-face class, respectively.

Discovering and noting that a difference exists between the two exams will not provide the insight that is needed. As the review of literature suggests, many studies find no difference between Blackboard courses and face-to-face ones while fewer studies actually find such difference. If a difference exists, factor analysis will be employed to uncover substantive factors that result in the different outcomes of the two tests. Factor analysis should provide a basis for establishing a theoretically based cause.

Identification of a cause provides two critically important contributions. First and foremost, it addresses type one errors (i.e., reducing the chance of thinking there are different effects while in fact the two teaching methods essentially yield similar student performances). Although remote, type one errors remain a problem. The existence of a type one error can be conceptually reduced (*not* statistically reduced) if a clear pattern within the difference emerges from the analysis. Secondly, if a pattern can be identified, future hypotheses can be formed based on the uncovered pattern. Future and further studies should be testing the specific substantive difference between Blackboard courses and face-to-face ones.

Most importantly, if differences in specific types of exam items emerge in future research, the format for online courses must be rethought. New and creative strategies would be necessary to assure that online students receive the same quality of instruction and learning opportunities as face-to-face students.

2. REVIEW OF LITERATURE

Much of the past research on distance education versus face-to-face revolves around both student and faculty perceptions of online instruction and the design and implementation of such instruction. Research into significant differences in test scores between the two modes of instruction is growing, but not abundant. Within the body of existing literature, several studies suggest no significant difference between online education and face-to-face instruction in terms of test scores, whether the online instruction is via Blackboard or some other method of delivery.

One of the earlier studies (Davies & Mendenhall, 1998), investigated a course in health and physical education. Although students preferred the classroom experience over the web course, a comparison of test scores showed no statistically significant differences. Later research of different types of courses at various locations in the United States supports this finding as well (Kennedy & McCallister, 2001; Moorhouse, 2001; Imig & Bailey, 2002; Neuhauser, 2002; Parker & Gemino, 2001; Petrrachi, et al., 2005). Kennedy & McCallister (2001) compared the effectiveness of traditional teaching, email teaching and hybrid classes for a graduate introductory statistics class in terms of test scores on multiple choice exams. Moorhouse (2001) compared midterm scores for two MBA courses— one online and one classroom. No significant differences between the two emerged. Similarly, Parker and Gemino (2001) found no significant difference in final exam scores over five semesters for a business administration course offered both online and in the classroom. Imig and Bailey's (2002) interpretation of quantitative measures (quiz scores, assignments and overall final point total) suggested no overall difference between internet classes and traditional ones. Neuhauser (2002) examined the learning effectiveness of online versus face-to-face instruction with sixty-two college students. As with the other researchers, her results did not reveal any significant difference between the two groups of test scores or final grades. In the area of social work, Petrrachi's, et al. (2005) study of social work practice course in both online and traditional formats suggest that students learn similarly regardless of format.

Some evidence exists to suggest that distance education students actually perform better on test scores (Allen, et al., 2004; Alstete & Beutell, 2004; Schachar & Neumann (2003); Schoenfeld-Tacher, McConnell, & Graham, 2001). Allen, et al. (2004) used a meta-data analysis to summarize the quantitative literature comparing the performance of students in distance education versus traditional classes. This analysis indicated that distance education course students slightly outperformed traditional students on exams and course grades. In a similar comparison, a meta-analysis of final course grades from 1990 to 2002, Schachar and Neumann (2003) state that, in two-thirds of the cases studied, students taking courses by distance education outperformed their counterparts enrolled in traditional courses. Schoenfeld-Tacher, McConnell, & Graham (2001) examined the effects of distance delivery in an upper histology course, and found that students in the online group outperformed their peers in the class setting. However, Alstete and Beutell (2004) note that students with "interest in online courses" do better than students who are in an online course, but would rather be in a face-to-face setting. This study found discussion board performance to be positively correlated with online performance in general.

On the other hand, some research indicates lower performance by students in online courses (Faul, et al. (2004); Faux & Black-Hughes (2000); Hisle-Gorman & Zuravin, 2006). Hisle-Gorman and Zuravin (2006) studied the performance of 73 MSW students in a social work research course. Their results indicated that students in either a traditional lecture course or a hybrid one had better midterm, final, and composite exam grades than those in a Blackboard only class. It was noted that professors teaching with Blackboard were doing so for the first time – a variable possibly having an effect. This research is consistent with Faux and Black-Hughes (2000), who found that social work history students in a traditional course gained more knowledge than those in an online format. In addition, Faul, et al. (2004) found higher performance from students in a lecture only course versus a hybrid one.

Much of the research comparing online instruction with traditional methods has focused specifically on test scores, regardless of the content of the tests. Applying concepts to practical situations is an area that is beginning to be explored. The application of learned skills to actual practice is a challenge for educators in many fields, including social work.

Ouellette, et al. (2006) compared interviewing skills in an undergraduate social work practice course, with one group being taught in a classroom setting and another in an online format with a different instructor. Results indicated no statistical difference between the two groups. However, Wachenheim (2004) discovered that, although online students did better on exams and homework, classroom students demonstrated a greater ability to apply course concepts to a practical setting. Wachenheim's results are interesting in light of research done by Weems (2002). This study compared two sections of beginning algebra – one taught online and one onsite. Although there was not a significant difference between exam averages for the two formats, there was a significant decrease in performance by the online students across the exams, whereas performance by the onsite students remained stable. Both the Wachenheim and Weems research suggest that a strict comparison of grades might not present the entire picture of student achievement between the two modes of instruction. In addition, they suggest possible differences in applying concepts to practical situations.

Wilke and Vinton (2006) provide support for the comparison research presented herein. In their study of the online MSW program at Florida State University, education and satisfaction outcomes are similar for the online and face-to-face groups – except for field work. When online students must apply practical concepts to real practice situation, they do not fare as well as students involved in a face-to-face educational experience. Others (Ouellette, et al., 2006 and Siebert, D.C. & Spaulding-Givens, J., 2006) have expressed concern about social work practice skills being taught online. Our study also suggests that online students may have difficulty applying concepts to practice.

While more evidence exists suggesting that online students can attain knowledge and skills as effectively as those in a classroom environment, the results are mixed, with questions arising regarding the content of a course compared to the delivery, as well as the application of theoretical knowledge to practice. A strict comparison of grades across the board may not provide a full picture of learning.

3. THE COURSE: *SOCIAL WORK VALUES AND ETHICS*

3.1. Overview

Social Work Values and Ethics is a simple one-hour course that reviews the NASW Code of Ethics. The syllabus for the face-to-face course can be found at:

http://www.uncp.edu/home/marson/Personal/Syllabi/450S.html

The syllabus for the Blackboard course can be found at:

http://www.uncp.edu/home/marson/Personal/Syllabi/450_online.htm

In each course, the NASW Code of Ethics is reviewed and implications for BSW practice for each standard are addressed in two manners. First, the language of the standard is restated in terms that would be meaningful for BSW practice (avoiding MSW practice issues). Second, a specific example of how the standard is applied in BSW practice is introduced. In the face-to-face section, the standard and discussion are completed orally. On Blackboard, each standard and discourse is found within the "Discussion" boards. The examples for both classes are the same.

However, several differences can be noted. Students in the face-to-face class are not required to comment or discuss each standard that is introduced. Students in the Blackboard class are required to participate in every discussion. In addition, if a student in the face-to-face course misses a class, he/she misses the discussion. Students in the

Blackboard course participate in an asynchronous manner.

As can be seen on the two syllabi, the content of the two sections is identical. The delivery of the content was the major difference. A common final exam was shared with both sections.

3.2. The Final Exam

The final exam was based on the content of the National Association of Social Workers *Code of Ethics.* Ten essay questions constituted the exam. These questions can be divided into three categories:

1) basic knowledge of code;

2) application of an abstract ethical concept in a practice scenario;

3) both knowledge and application of code.

The final exam emphasized ethical issues that were most likely to occur to BSWs in entry level practice. For example, no private practice issues were included (i.e., billing), but dual relationships were emphasized because these relationships are commonly found in rural settings. Because the test items are included in a bank of essay items that may be used in the future, the specific items will not be presented.

Table 1 illustrates the content of the common exam for the Blackboard and face-to-face courses. The first column includes the order of the test items as administered to both Blackboard and face-to-face students. The second column identifies the type of item as defined in the preceding paragraph. The number in parenthesis [i.e., (2)] indicates the number of items for each type.

Thus, "Abstract to Practice (2)" would indicate the second item on the exam that addresses the application of an abstract ethical concept in a practice scenario. Column three identifies the specific code that was the focus of the item. The entire *Code of Ethics* with the corresponding Standard number can be found at:

http://www.socialworkers.org/pubs/code/code.asp

Test item number 5 is a general overview of the entire *NASW Code of Ethics*. This table enables the reader to see the specific code that was the heart of each test item.

Table 1

Table 1.   Link of Ethical Code and Item Content

Test Item	Ethical Code	Ethical Content of the Item
1	Abstract to Practice (1)	Standard 1.06c: Conflict of Interests, dual relations
2	Abstract to Practice (2)	Standard 1.06c: Conflict of Interests, dual relations
3	Ethical Code (1)	Standard 2.10a and b: Incompetence of Colleagues
4	Abstract to Practice (3)	Standard 1.03a, b, c, d, e, and f: Informed Consent
5	Both (1)	Standard all of them
6	Ethical Code (2)	Standard 1.06c: Conflict of Interests, dual relations
7	Ethical Code (3)	Standard 4.01a and b: Competent Service Limitations, Competence in New Techniques
8	Abstract to Practice (4)	Standard 1.05a and 1.10: Cultural Competence and Social Diversity, Physical Contact
9	Ethical Code (4)	Standard 1.05a, 1.10, 2.10b: Cultural Competence and Social Diversity, Physical Contact, Incompetence of Colleagues
10	Both (2)	Standard 2.07b: Sexual Relationships

4.3. STATISTICAL ANALYSIS

4.1. T-Test

For all the 23 students who enrolled in the blackboard course, their mean total scores were calculated. Similarly, for all the 16 students enrolled in the face-to-face class, their mean total scores were computed. Before the comparison of the mean total scores between these two groups of students is performed, an F-test is conducted to determine whether or not the two data sets are of equal variance. The F-value is 1.07, resulting in the probability $P(F > 1.07) = 0.9$ (folded F; degrees of freedoms are 22 for numerator and 15 for denominator). This means that the two data sets should be considered as equal variance. Accordingly, the T-test for equal mean uses the pooled variance, resulting in a probability $P(T > 2.59) = 0.014$ (degrees of freedom is 37) indicating that the average scores between face-to-face and Blackboard classes are significantly different at the significance level $\alpha = 0.014$. These results of the T-test are shown in Table 2 below.

Since the mean score of 8.87 for the face-to-face class is higher than the mean score of 8.14 for the Blackboard class, the conclusion from the T-test is that the overall response from face-to-face students is

better than that from blackboard students. Moreover, the 95% confidence interval of the mean score for the face-to-face class is [8.42, 9.32] with a standard error of 0.21; it is [7.77, 8.52] with a standard error of 0.18 for the blackboard class.

Table 2

Table 2. Comparison of Face-to-Face vs. Blackboard Mean Test Scores

		Mean			Std dev			Std Err
Group	N	Lower CL	Mean	Upper CL	Lower CL	Mean	Upper CL	
Face-To-Face	16	8.42	8.87	9.32	0.62	0.84	1.31	0.21
Blackboard	23	7.77	8.14	8.52	0.68	0.87	1.24	0.18
Difference		0.16	0.73	1.29	0.70	0.86	1.11	0.28

T-test Results

Method	Variables	DF	t Value	Pr > |t|
Pooled	Equal	37	2.59	0.014

4.2. Factor Analysis[1]

Based upon the calculated eigenvalues of the correlation matrix between the ten variables for each group, the first four principal components (primary factors) are selected for each group to allow a comparison between the two groups, which is based on the Henry Kaiser's Eigenvalue-Based Rule (Rule of Thumb): The number of factors is chosen as the number of eigenvalues of the correlation matrix that are larger than 1 (Morrison, 1990). The principal components (PCs) are extracted so that the first component accounts for the largest amount of the total variation in the data, i.e., the first component $PC_{(1)}$ is that linear combination of the observed variables X_k (k = 1, 2, …, 10)

$$PC_{(1)} = w_{1,1} X_1 + w_{1,2} X_2 + … + w_{1,10} X_{10} \quad (1)$$

where the weights $w_{1,1}$, $w_{1,2}$, … , $w_{1,10}$ will be chosen to maximize the ratio of the variance of $PC_{(1)}$ to the total variation existing in all X_k, subject to the constraint $\sum w^2_{1,j}$ = largest eigenvalue (the sum is over all k's from 1 to 10). The second principal component $PC_{(2)}$ is that weighted linear combination of the observed variables which is unrelated with the first linear combination and which accounts for the maximum amount of the remaining total variation not already accounted for by $PC_{(1)}$, where the sum of all

squared weights is equal to the second largest eigenvalue. In general, the mth principal component is that weighted combination of the X_k's which has the largest variation of all linear combinations that are unrelated with all of the

$$PC_{(m)} = w_{m,1} X_1 + w_{m,2} X_2 + … + w_{m,10} X_{10} \quad (2)$$

previously extracted principal components, where the sum of all squared weights is equal to the mth largest eigenvalue.

Factor analysis can simultaneously manage over a large number of variables, compensate for random error and invalidity, and disentangle complex interrelationships into their major and distinct regularities.

Factor analysis attempts to simplify complex and diverse relationships that exist among a set of variables by uncovering common dimensions or factors that link together the seemingly unrelated variables, and consequently provides insight into the underlying structure of the data. In other words, factor analysis reduces variables into a smaller set of factors which explain the variance in the original variables.

Since there is not an available theoretical hypothesis, the factor analysis conducted here is exploratory, not confirmatory (when some prior information on the common structure underlying the data is given and one wishes to confirm or negate the hypothesized structure). The results of Rotation Method Promax are summarized in Table 3 below for face-to-face and blackboard data respectively.

4.3. Factor Analysis: Face-To-Face Data

From the SAS outputs in Table 3 below, the first four clusters (i.e. factors) account for 80.5% of the total variation in the scores. Factor 1 (comprised most strongly of items 6 and 8 through 10) alone accounts for an approximate 29.7% of the total variation, and a larger variation in scores for these students appear in

this factor. Factor 1 primarily involves the theory of the ethical codes regarding relations and culture.

Factor 2 shows high loadings for items 2, 5 (negative loading; refer to the explanation given in Section 5) and 7. This factor primarily contains ethnical competence. Noting that the loading for item 7 is stronger here than it was for Factor 1, we drop it from the interpretation of Factor 1 and use it, along with items 2 and 5, to interpret Factor 2. Factors 3 and 4 can be interpreted in a similar way.

Variations explained by these clusters are given in the last row of Table 3. This trend of performance is normal. Moreover, the SAS outputs from the Principal Component, Varimax and Promax are quite consistent for the data.

Table 3

Question	Topic	Face-to-face class				Blackboard class			
		Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4
1	Abstract to Practice (1)	-0.22	-0.13	0.00	**0.88**	**0.59**	**0.61**	0.04	0.49
2	Abstract to Practice (2)	0.13	**0.78**	0.29	0.23	0.29	**0.60**	-0.18	**0.66**
3	Ethical Code (1)	-0.04	0.17	**0.90**	0.01	0.06	0.06	**0.80**	0.08
4	Abstract to Practice (3)	0.16	-0.16	**0.87**	-0.03	0.14	-0.07	0.28	**0.87**
5	Code/Abstract to Practice (1)	0.02	**-0.69**	0.12	0.24	-0.03	**0.80**	0.16	-0.04
6	Ethical Code (2)	**0.62**	0.03	-0.06	**0.68**	**0.88**	-0.03	-0.04	0.23
7	Ethical Code (3)	0.40	**0.78**	-0.12	-0.15	**0.56**	0.36	-0.07	-0.09
8	Abstract to Practice (4)	**0.83**	-0.10	0.37	0.00	**0.67**	0.32	0.20	0.35
9	Ethical Code (4)	**0.87**	0.27	-0.02	-0.10	0.35	**0.56**	-0.06	0.12
10	Code/Abstract to Practice (2)	**0.95**	0.24	-0.06	-0.01	0.04	0.04	**0.83**	0.24
	Variance Explained	2.97	1.90	1.81	1.37	2.13	1.93	1.51	1.68

3.4. Factor Analysis: Blackboard Data

An increased number of blackboard students could not completely grasp and handle more complex concepts and apply them.

From the SAS outputs in Table 3, the first four factors account for 72.5% of the total variation in the scores compared to 80.5% for face-to-face class. Factor 1 (comprised mostly of Items 1 and 6 through

8) alone accounts for only 21.3% of the total variation compared to 29.7% for the face-to-face class. Factor 1 primarily involves ethical competence. Noting that the loadings for items 9 and 10 are significantly lower than that for face-to-face class, implying that less variation exists on these items for blackboard class, i.e., for more comprehensive ethical codes and abstract to practice, most students of the blackboard class did not perform well (for face-to-fact class, however, some performed much better

than others). Hence, students have experienced more challenges regarding concepts and applications.

Factor 2 shows high loadings for items 1, 2, 5 (positive loading), and 9. Noting that the loading for item 1 is slightly higher here than it was for Factor 1, we used it along with other items to interpret both Factors 1 and 2. Factor 2 contains primarily relations and culture. For Factors 3 and 4, the discussions are similar.

This trend of performance for the blackboard class is abnormal, implying additional communication methods are needed to enhance distance education and learning. In addition, the SAS outputs from the Principal Component, Varimax and Promax are quite different for the data.

4. RESULTS

In summary, the F-test and t-test demonstrate a statistically significant difference in the outcomes of the final exam for the Blackboard and face-to-face students: While the F-test indicates that the variances of the two sets of scores are close, the pooled t-test indicates that the average score of the face-to-face students is better than that of Blackboard students with a small P-value ($P = 0.014$). Hence, the overall performance of face-to-face students for the final exam is better than the Blackboard students.

In further elaboration, the results of individual factor analyses for the two sets of scores suggest two different patterns of performance. Within the face-to-face class, students have similar performance resulting in less variation for the first five questions but the variances are widely distributed for the last five questions as shown in Table 5, under Factor 1 of Face-To-Face class. In contrast, within the Blackboard Class, the variations in performance relative to all the test scores are distributed for questions 1 and 6 through 9, as shown in Table 5, under Factor 1 of Blackboard Class. Notice that these variations are not the variations calculated from individual questions as they are also related to the grand mean. Another concern is the negative loadings. While factor loadings obtained from a rotation often give useful interpretations, negative loadings cannot be interpreted as correlation coefficients. In such a case, an appropriate interpretation is that variables with positive loadings and those with negative loadings contribute to the performance measure in the reversed ways.

Finally, for the face-to-face class, the first and second factors are relations and culture, and ethnical competence respectively. For the blackboard class, there are reversed to ethnical competence, and

relations and culture. Moreover, for the factor loading on Item 5 that involves both the theory of all ethnical codes and the applications to practice, the loading is positive for the face-to-face class, but negative for Blackboard class.

5. STRENGTHS AND LIMITATIONS OF STUDY

Strengths and limitations of comparison studies are addressed by the amount of control the study has for alternative explanations. Within this study, we must address the issue of internal validity by the question, "Is Blackboard the cause of the difference between the scores on the final exam OR could an alternative phenomenon have caused the difference?" Part of the answer to this question rests within the statistical analysis. In addition, one must consider issues of control or the elimination of alternative explanations. Below are listed the characteristics of the study that demonstrate control followed by characteristics that illustrate limitations of control.

5.1. Strengths

- The students were very similar. All were juniors and seniors social work majors who had completed the same course prerequisites.

- The final exams were identical.

- Although the sections of the course were not offered during the same semester, the Blackboard course (the one with the lower scores) was offered second. This means that the Blackboard students had an advantage over the face-to-face students. If the Blackboard student queried the face-to-face students about the questions on the exam, they would have had the exam items in advance. Since their scores were lower, there is no indication that the Blackboard students made such an effort.

- All course assignments were identical.

- Both exams were given on Blackboard for each class.

- Both final exams were based on the identical course assignments.

- There were no significant differences in course assignment grades between the two sections.

5.2. Limitations

- The study can best be described as *ex post facto*. There was no thought of comparing the results of these exams until the professor became disappointed with the performance of the Blackboard class. In this lack of advanced planning lies the fertile soil for a Type I error.

- One course was offered in the Fall of 2004 while the other course was offered in the Spring of 2005.

- The assignments for the Blackboard course were given in writing, while the assignments for the face-to-face course were given orally (but were also found on the syllabus).

- Blackboard students submitted their assignments electronically while the face-to-face students completed their assignments orally during class time.

- No random selection was employed for the two courses. In addition, students admitted to the Blackboard course were required to have Blackboard experience.

6. Conclusion and Recommendations for Further Study

As noted in the review of literature, most studies find little or no difference between knowledge obtained in face-to-face and Blackboard classes. However, in many ways, virtually all this literature is broad and/or cursory (DeNeui & Dodge, 2006). One major exception is the study produced by Wilke and Vinton (2006). They also suggest that Blackboard is weaker with the application of concepts to real-life social work practice. Our study's primary objective was to lay out specific areas for further investigation. Replication of this study with much greater experimental control is necessary. The new research question is: "Are face-to-face students better equipped to understand and apply abstract and theoretical concepts in social work practice than Blackboard students?"

Based on what we learned from the current research, the ideal research conditions to advance our knowledge of Blackboard would include the following experimental controls:

- There should be two sections of the same course in which students are randomly assigned to Blackboard or face-to-face courses.

- A single professor should teach both sections.

- Identical assignments should be given to both classes.

- Identical exams are given to both classes.

- The content of the exam must be limited to test items that focus on abstract concepts that are applied to social work practice.

- Test items should be both essay and multiple choice.

- Both exams must be administered on Blackboard.

- Application of concepts must be the central focus of all test items. Although reasoning/problems solving and recall are considered important aspect of learning, past research does not support a difference with these categories of test items. Thus, test items that focus on reasoning/problem solving and recall would contaminate the statistical inference that could be made. [Note: Test items that focus on application of concepts are the most difficult to compose.]

- A single professor must grade essay items, while multiple choice items can be graded on Blackboard.

- The grading professor must be blind to knowing the author of the essay exam and the class in which the student was enrolled.

With replication employing these controls, we will gain greater insight into the strengths and weaknesses of Blackboard. If there continues to be a difference between Blackboard and face-to-face, online courses in social work need to be reexamined. This is not to suggest that Blackboard might be an inferior protocol for teaching, rather that future research could direct the academic world to new strategies for teaching a Blackboard class.

7. Some Thoughts on the Project

The data for this study was collected during the academic year 2004-2005. During the spring semester of 2007, I (SMM) once again taught the SWK 4500 Social Work Values and Ethics as a Blackboard course. During the fall, I taught the same course, but face-to-face. Within the Blackboard course, I had only 7 students. I knew all of them from past courses.

Within the rural practice arena, the introduction of strategies for addressing unavoidable dual relationships is critical, but difficult to understand. For Blackboard, I include a film and graphic to illustrate how best to handle unavoidable dual relationships. This is the same graphic and presentation provided to students in the face-to-face class. As part of the Blackboard discussion, I restated the strategy within the discussion board but invited the students to meet me to resolve any confusion. One of the students came to my office and asked questions. During the final, the same student was the only one who gave the correct answer for addressing unavoidable dual relationships.

Although the evidence is somewhat weak, I firmly believe that students enrolled in Blackboard courses have trouble applying abstract concepts to practice. I believe that our state-of-the-art measurement strategies are not sophisticated enough to clearly delineate these differences. We have strong administrative advocates for Blackboard because this online delivery system is economically profitable. Blackboard and other online strategies are here to stay. It is incumbent on all social work faculty to look deeper into the implications of using Blackboard.

8. References

Allen, M., Mabry, E., Mattrey, M., Bourhis, J., Titsworth, S., and Burrell, N. (2004). Evaluating the effectiveness of distance learning: A comparison using meta-analysis. *Journal of Communication*, *54(3)*, 402-420.

Alstere, J. W., & Beutell, N. J. (2004). Performance indicators in online distance learning courses: A study of management education. *Quality Assurance in Education, 12*(1), 6-14.

Anderson, J. W. (2003). Faculty perspectives of the Blackboard course delivery system. Presented at the Annual Meeting of the Mid-South Educational Research Association, November 5-7 (Biloxi, MS).

Davies, R. S. & Mendenhall, R. (1998). Evaluation comparison of online and classroom instruction for HERE 129-Fitness and Lifestyle Management course. (ED 427 752).

DeNeui, D. L. & Dodge, T. L. (2006). Asynchronous learning networks and student outcomes: The utility of online learning components in hybrid courses. *Journal of Instructional Psychology, 33*(4), 256-259.

Faul, A., Frey, A., & Barber, R. (2004). The effects of web-assisted instruction in a social work methods research course. *Social Work Education*, *23*(1), 105-118.

Faux, T., & Black Hughes, C. (2000). A comparison of using the internet versus lectures to teach social work history. *Research on Social Work Practice, 10*(4), 454-466.

Fritz, K. M. (2003). Using Blackboard 5 to deliver both traditional and multimedia quizzes on-line for the foreign language class. Internal Report for the University of South Carolina.

Gabbard, D. A., L'Esperance, M., Perez, T., & Atkinson, T. (2002). Dealing with disengagement through diversity. *Multicultural Education, 10*(2), 41-48.

Hisle-Gorman, E., & Zuravin, S. (2006). Teaching social work research: A comparison of web-based and in-class lecture methods. *Journal of Technology in Human Services, 24*(4), 77-92.

Imig, D.R. and Bailey, D.C. (2002). *Traditional* v*ersus* internet instruction: A comparative assessment of *student* learning. *Journal of Teaching in Marriage and Family, 2*(1), 69-82.

Kennedy, R. & McCallister, C. J. (2001). Graduate introductory statistics in class vs. online. Presented at the Annual Meeting of the Mid-South Educational Research Association (30th, Little Rock, AR, November 14-16).

Mariola, E., & Manley, J. (2002). Teaching finance concepts in a distance learning environment—A personal note. *Journal of Education for Business, 1*(1) 177-180.

Moorhouse, D. R. (2001). Effect of instructional delivery method on student achievement in a master's of business administration course at the Wayne Huizenga School of Business and Entrepreneurship.

Morrison D. F. (1990). *Multivariate Statistical Methods*. McGraw-Hill, New York.

Neuhauser, C. (2002). Learning style and effectiveness of online and face-to-face instruction. *American Journal of Distance Education, 16*(2). 99-113.

Ouellette, P.M., Westhuis, D., Marshall, E., and Chang, V. (2006). The acquisition of social work interviewing skills in a web-based and classroom instructional environment: results of a study. *Journal of Technology in Human Services, 24*(4), 53-75.

Parker, D., & Gemino, A. (2001). Inside online learning: Comparing conceptual and technique learning performance in place-based and ALN formats. *Journal of Asynchronous Learning Networks, 15* [http://www.aln.org/publications/jaln/v5n2/index.asp]

Petracchi, H., Mallinger, G., Engel, R., Rishel, C. W., and Washburn, C. Evaluating the efficacy of

traditional and web-assisted instruction in an undergraduate social work practice class. *Journal of Technology in Human Services 23*(3/4), 299-310.

Shachar, M., Neumann, Y. (2003), "Differences between traditional and distance education academic performances: a meta-analytic approach", *International Review of Research in Open and Distance Education*, *4(2).* Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/153

Schoenfeld-Tacher, R. McConnell, S., & Graham, M. (2001). Do No Harm—A Comparison of the Effects of On-Line vs. Traditional Delivery Media on a Science Course. *Journal of Science Education and Technology, 10*(3), 257-65.

Siebert, D.C., and Spaulding-Givens, J. (2006). Teaching clinical social work skills entirely online: A case example. *Social Work Education 25*(1), 78-91.

Topper, A. (2005). Facilitating student interactions through discursive moves: An instructor's experience teaching online graduate courses in educational technology. *Quarterly Review of Distance Education, 6*(1) 55-67.

Wachenheim, C. J. (2004). How a online course compares to its classroom counterpart: A preliminary investigation. *NACTA Journal, 48*(4) 20-26.

Weems, G. H. (2002). Comparison of beginning Algebra taught onsite versus online. *Journal of Developmental Education 26*(1), 10-12.

Wilke, D., & Vinton, L. (2006). Evaluation of the first web-based advanced standing MSW program. *Journal of Social Work Education, 42*(3) 607-612.

---

[1] For factor analysis, it is preferred that the data size be 5 to 10 times the number of questions.

The book titled, *Women of Color on the Rise: Leadership and Administration in Social Work Education and the Academy* (2010) is a compilation of narrative essays generated from a diverse group of women with experience in leadership and administrative roles within academia. Halaevalu F. Ofahengaue Vakalahi and Wilma Peebles-Wilkins served as the editors of this narrative work. Vakalahi has a history of fulfilling leadership and administration positions at various academic institutions and is currently a director of the masters level social work program at George Mason University. Peebles-Wilkins has been an administrator in public, private, and research settings. Furthermore, she is noted as an NASW Social Work Pioneer and is currently fulfilling the role of dean emeriti of the Boston University School of Social Work. The editors have sectioned the book into four parts that highlight the stories of women in various leadership and administrative positions including deanship and chair/directorship. To enhance the level of diversity within this writing, each essay contributor has yielded from one of the five racial/ethnic populations "federally defined" as "historically disadvantaged" (p. xi).

Women of Color on the Rise communicates the contributors' devotion to the NASW Code of Ethics (1999) that is perpetuated through their positions of leadership. This writing gives credit to the accomplishments of persons in the past and present and provides both guidance and inspiration for those interested in academic leadership and administration. Each contributor shared her story about leadership within the framework of 13 questions. This framework allowed for themes including the influence of families and mentors, specific strategies for professional growth, coping methods to address challenges, and guidance for future generations of leaders.

This writing has depth in communicating a strengths-based approach to self-evaluation, the impact of various systems on the career growth of individuals, in addition to the use of resilience, motivation, and application of positive coping strategies for goal achievement. As a result, this book may be used to enhance diversity, management, and indirect practice social work course content. In addition, the essays can be applied as a mentorship tool to influence student understanding of administrative roles in social work and encourage future growth in leadership competency. Furthermore, as this writing can help promote leadership development, it can also be a means to provide insight for the educator. As a result, educators may develop an increased understanding of what undergraduate social work students may be experiencing regarding familial influence, integration challenges into the academic environment, and the level of interest in mentorship.

Overall, the concepts expressed through this writing may be applied toward recruitment and retention efforts of both students and professionals in the academic arena. As *Women of Color on the Rise* highlights the realities of barriers to acceptance, challenges to self-esteem, and periods of isolation, the lessons of these stories provide action steps toward addressing those challenges and pursuing leadership development. As a recognizable aim of this book is to further promote "academic systems and institutions" (p. ix) that respond to the concepts of diversity, cultural strengths, as well as social change and justice, it may productively contribute to the further enhancement of the social work profession.

Dr. Lieberman is a Chancellor's Club Teaching Professor of Social Welfare at the University of Kansas. She has published multiple books focusing mainly on effective social work practices. Her work emphasizes understanding diversity and building upon women's strengths. Dr. Lieberman has also collaborated on several federally funded family and children projects.

This book profiles 15 women social workers who have succeeded in contributing to profound changes in their communities and countries around the world. Although the author does not specify a particular audience, I believe this book is as suitable for a general audience as it is for an academic audience. Whether a person is simply interested in social justice or is trying to learn what practices are successful in diverse environments, this book is a good choice.

The book is organized into three parts. The first part focuses on five social workers who have gained positions that enabled them to create and change national and local policies. The second part features women who have created social change from the bottom up through community organizing. Part III presents the biographies of three women who have used their social work knowledge to bring great change to repressive and developing countries.

Although these women's social class, religious, and cultural backgrounds are diverse, their desire for social justice and the core values learned in their social work education connect them. All of the women claimed that their social work educations provided them with the knowledge of structure and application that prepared them to accomplish their goals. Each of the featured women overcame structural or personal obstacles to achieve their accomplishments. Some came from impoverished backgrounds; others worked in countries where women have few rights. All of them give credit for their stamina and courage to mentors. Many cited their mothers as mentors. Additionally, many credited inspiring teachers.

This book has multiple strengths. The writing style is clear, concise, and accessible to the general reader as well as the social work scholar. Although the biographies are short, they provide poignant and detailed accounts of various obstacles faced and successes earned by each woman. This allows readers to see "social work in action." In addition, the biographies clearly show that the women did not have a set "plan of action" for their work but remained open and flexible so they could afford themselves of opportunities. The results of their willingness to change courses when needed can inspire all of us who seek to improve our communities. While I really enjoyed this book, I believe it would have been improved by the addition of a final chapter to tie the book together. As it is, the book just stops with the last biography. It would help students and general readers, alike, to summarize how what they have learned could benefit them in their own lives or careers.

Overall, I would recommend this book and believe it can serve multiple purposes. I believe it would be a good addition to any social work practice course. In addition, this book would make an inspiring addition to gender studies courses. And finally, I believe that this book could be used successfully in political science courses to illustrate how the core values of social work can help craft good social policy.

I was asked to review *Queer Political Performance and Protest* for two reasons. I have a long history of writing book reviews and am the Senior Editor for *The Journal of Social Work Values and Ethics*. In addition, I have taught a course entitled SWK 3870 Sexual Minorities. These factors can easily be Googled. I am a bit out of my element, but I am glad I had an opportunity to read this fine book.

Although Shepard does not explicitly identify an intended audience, it is clear that his work is intended for those who have an interest in the gay liberation movement. The content includes issues that might be of interest to lesbians, bisexuals, and transsexuals. However, the major theme of the book clearly is focused on the gay male population. In addition, to the gay liberation movement, those who are interested in community organization models – particularly community change – will find Shepard's work worthy of reading. Besides teaching Sexual Minorities, I teach Community Organization. I have already referred to *Queer Political Performance and Protest* as a good example of efforts and strategies for community change.

There are several points that I find necessary to address:

First, this is not a book for those who are unknowledgeable about the gay, lesbian, bisexual, transgendered social/political/historical civil rights movements. Shepard operates under the assumption that the reader is acquainted with the movements. While reviewing books, I often make margin notes. One of my notes read: "historical events can be almost meaningless if you weren't there." At times, it seems like Shepard is story-telling to a close friend who experienced the events. In that sense, his writing style is highly personalized and has a rather touching approach to explaining events and strategies. One normally does not read community change monographs with such a vision.

Second, as stated earlier, the community change and community organization examples are vivid and concrete rather than being abstract that are most commonly found in such macro monographs. These examples engaged micro social work students. In our current political climate, I see a stress on political philosophy (liberals vs. conservatives) and a slow death of American Pragmatism. Politicians (of both parties) are more interests in winning than in pursing change that has a chance of achieving a predetermined goal. That being said, the focus of Shepard's thesis thrust is highly pragmatic. This is a refreshing change in the recent literature that I have been reading. A good example of the American Pragmatism emphasis can be found in the efforts made in the needle exchange program starting in page 146. The strategies and insight can be generalized to other efforts in the arena of change and community organization.

Third, Shepard offers a manual in the use of humor in street conflict. He correctly points out that humor disarms the opposition (i.e., the police). Authorities simply do not know how to handle protesters who are laughing and joking. In addition to the confusion humor generates, it also reduces violence. It is very difficult for a police officer to beat a protester with a night stick if the protester is doing something considered to be funny. Shepard's analysis of humor is confirmed in McNamara's book, *In Retrospect*. McNamara notes that his staff were totally lost and had no strategy to control Vietnam protesters until they became violent [see page 303].

In the end, Shepard composed an appealing monograph. It would be of great interest to those involved in the gay liberation movement. Social work students, professors and community organizers will find it useful in generalizing to other issues related to social change and community organization. I recommend this book to be adopted by libraries that have holdings for sex research, social work, and community organization.

My background as a social worker is a collage of micro and macro practice experiences. I have worked in the fields of domestic violence, child welfare, and hospice. While working for a nationwide hospice, I progressed from a "front line" clinical social worker to a National Director of Hospice Development. Regardless of the setting in which I worked, I never encountered a neatly packaged problem or issue and there were no "textbook" cases with prepared solutions.

As a new professor in social work, I strive to have my practice courses reflect the day-to-day experiences of social workers in the field. Thomas O'Hare's *Essential Skills of Social Work Practice* provides students with knowledge regarding basic social work skills, but more importantly, he provides case studies in which he skillfully identifies each and every aspect of the case from both the patient and the practitioner's perspective. It is in these case studies, that students experience the application of the skill sets that they will need each and everyday they work in the field.

The text opens with a clear overview instructing the reader on the connection between assessment, intervention and evaluation. It is the following two chapters that O'Hare distances his text from other introductory practice books, in these two chapters he writes on the connection between research and practice and ethics in social work practice. I have not encountered many texts that give both of these topics the same prominence that is found in this book. He communicates the importance of these issues to the reader prior to discussion the "essential practice skills". Typically, in my experience, topics such as research and ethics are discussed as separate but important issues in practice.

Chapters four through seven outline, in exceptional detail, skills such as conducting an assessment, supportive skills, therapeutic coping skills and case management skills. As with Part I of the book, it is in his detail that O'Hare distinguishes himself and his attention to topics that other authors ignore at worst or minimally mention at best. His chapter on case management skills outlines the essence of casework and, at the same time, uses cases from various social work fields to illustrate his points to the reader.

Another unique direction taken by O'Hare is his delving into several disorder classifications and for each disorder grouping he takes the reader from assessment to intervention to evaluation. He uses case examples for each of these specific discussions. It is in his attention to detail and the methodical way in which he conducts the cases that illustrate the application of each of the concepts previously discussed in the book.

He concludes the book with a chapter on evidence-based practice. I was hoping that he applied the same level of rigor and analysis to this subject as his previous topics, however, I find this chapter somewhat lacking. O'Hare identifies four major objections to conducting evidence-based practices but does so without a great amount of detail or specificity. He then transitions to broader service delivery issues concluding the chapter soon thereafter.

Finally, the appendices of this book include a "psychosocial intervention scale" and a "comprehensive service plan". These tools are indispensible in social work practice, tools that are used to bring clarity to the at times overwhelming nature of the work that we do. In my opinion, this is an excellent text for a upper level undergraduate practice course and/or for a foundational practice course in a MSW program.

Betsy Murphy is a family nurse practitioner and is certified as a hospice and palliative care nurse with 23 years of experience in hospice care. Most recently, Ms. Murphy has focused on providing education through publication and presentation on hospice care. Ms. Murphy self-published *Guide to Caregiving in the Final Moments of Life* to help caregivers recognize the signs of imminent death. Ms. Murphy suggests that is it through such awareness that caregivers can avoid denial, advocate for, and prepare with those facing the end-of-life.

This booklet begins with a review of the early signs of death and complications associated with the dying process. Early signs of death include poor appetite, weight loss, weakness, and dependency. This trend may not be reversed since patients have a compromised immune system that increases one's risk for infection. The dying process can thus lead to pneumonia, sepsis, heart trouble, and organ failure. At the end of life, additional symptoms include fatigue, changes in breathing patterns, refusal of food and drink, chronic pain, confusion, incontinence, encopresis, and restlessness. This booklet concludes with a brief discussion of common reactions to the death of a loved one as well as the role of hospice and palliative care.

Relative to social work values and ethics, awareness of such symptoms carries with it the responsibility of informing family members that the end is near. This information allows family members time to prepare as well as determine if heroic measures such as the use of e.g., a feeding tube, kidney dialysis, ventilator, or hospitalization are appropriate at the end of life. One of the most helpful parts of this booklet includes caregiving tips for each symptom of the dying process. For example, lip balm may be applied to the patient's mouth to reduce cracking given dehydration. Soft foods are recommended since they are easier to swallow when patients are too weak to chew. Rotating a patient in the bed with a draw sheet positioned under the patient can prevent skin break down. Messaging lotion on a patient's skin and just being present are identified as additional sources of comfort.

The text is written in a way that any reader could comprehend the material. Hence, this material is best suited for the general education of family members and nonmedical professionals involved in hospice and palliative care. However, editorial issues result in repetition, inconsistent formatting, and limited focus in the text. For example, there is more information on the dying experience of the elderly with dementia rather than symptoms associated with other diseases (e.g., congestive heart failure) (Levenson, McCarthy, Lynn, Davis, and Phillips, 2000) and populations (e.g., children) (Morrow, 2009). There is also little reference to academic literature and other resources for follow-up unlike a similar booklet produced by the National Institute on Aging (2008).

Some of the basic information included in Murphy's booklet can be found on the Internet through magazine and association publications as well as health education websites (see FMER, 2009; Hospice Patients Alliance, n.d.; Lamers, 2009; NIA, 2008). What makes this booklet distinct, though, is the collection of material in a form that makes it easy to distribute. The low cost of this item ($5.00 per booklet) further makes this an affordable resource. The booklet also focuses on what caregivers can do in response to the symptoms of pending death,

which can facilitate greater collaboration with health care professionals. Therefore, I would recommend *Guide to Caregiving in the Final Months of Life* for patient and practitioner education.

References

Foundation for Medical Education and Research (FMER) (2009). *End of life: Caring for your dying loved one.* Retrieved from http://www.mayoclinic.com/health/cancer/CA00048

Hospice Patients Alliance (n.d.). *Signs and symptoms of approaching death.* Retrieved from http://www.hospicepatients.org/hospic60.html

Lamers, W. (2009). *Signs of approaching death.* Retrieved from http://www.hospicefoundation.org/endOfLifeInfo/signs.asp

Levenson, J.W., McCarthy, E.P., Lynn, J., Davis, R.B., and Phillips, R.S. (2000). The last six months of life for patients with congestive heart failure. *Journal of American Geriatric Society, 48*(5 Suppl), 101-109.

Morrow, A. (2009, January 28). [Review of the book *Guide to caregiving in the final months of life*]. *About.com Palliative Care*. Retrieved from http://dying.about.com/od/reviews/gr/guide_to_care.htm?p=1

National Institute on Aging (NIA) (2008). End of life: Helping with comfort care. Retrieved from www.nia.nih.gov/NR/rdonlyres/.../End_of_Life_booklet_final.pdf

Reviewed by Ashley Bunnell Leonard, BSW, and Stephen M. Marson, Ph.D., University of North Carolina at Pembroke

This review is written with two voices: One (Ashley) who has a specialty in Child Protective Services (CPS), while the other (Steve) has a specialty in gerontology. This book will generate interest in both areas of social work practice and research. Within the Foreword, Dr. Andrew Cave (psychiatrist) states that the intended audience includes physicians and patients with Multiple Personality Disorder and Dissociative Disorder. Later, he writes that primary health care disciplines, practitioners, and students would find the book useful. We feel that this book would have a strong appeal to CPS workers (both BSW and MSW) and social work professors who teach child welfare courses. In fact, we agree that *If You Tell … It Will Kill Your Mother* should be an excellent required reading in a child welfare course. In addition, we believe that those who specialize in gerontology will find that this book answers some important aging questions – to be addressed later.

If one would have to categorize *If You Tell … It Will Kill Your Mother,* the book would best be characterized as a quasi autobiography. Most of the chapters are written in the first person by Ardith Trudzik. She unfolds a story of her life as a victim of vicious sexual abuse and exploitation. From her experiences emerge an array of psychiatric diagnoses. The repression of these sexual experiences appears to have exacerbated her emotional disorders which finally accumulated into a "nervous breakdown." The breakdown was the precipitating factor that lead to the acknowledgment that Ardith had gallantly suppressed her experience–at the expense of her emotional well-being. Simply stated, her ability to repress her experiences led to a worsening of her psychological state.

Each chapter shares a common writing pattern. The chapters are constructed with a unique style where

Ardith offers the reader the ability to engage in the content. Each chapter is written with various font changes to keep the reader connected with various color drawings and poems. At the end of each chapter, the author leaves the reader with the psychiatrist's analysis of the therapy sessions involving Ardith's life events. The linkage between Ardith's description of her life events and the therapist's analysis provides a profound level of insight. Social workers commonly employ the term *mutuality* to describe this dynamic.

There are several major aspects of this book that provide vital information which is particularly important for the education of BSW and MSW social work students. By reading the book, CPS practitioners will gain critical insight that they would not have otherwise received. These insights include but are not limited to the following:

- Information in the book will expand students' and practitioners' knowledge of Multiple Personality Disorder and Dissociative Disorders. These disorders are common outcomes of sexual child abuse.
- Ardith's description of her world view offers great insight to the world of sex abuse in a manner that could *never* be achieved by intensely studying diagnostic manuals, research and other material usually required for academic assignments and in-service training.
- The detection of child sexual abuse is an emotionally complex enterprise. Ardith opens the private door to the world of sexual child abuse in a profound manner. She offers the social worker insight into how to articulate statements and questions that would be most effective in eliciting best

responses from a victim without inducing undo emotional distress.

- The old adage "begin where the client is" emerges as a common theme. The trusting relationship between the practitioner and the victim must be paramount in all intervention strategies. This basic trust is the vital catalyst for successful treatment. Treatment cannot be hurried.
- An understanding of basic medical terms is essential for communication between social work practitioners and physicians who are part of the intervention team. When a social work practitioner lacks a basic understanding of human biology, treatment can be stifled.
- Victims of sexual abuse *withhold* information. In Ardith's case, she withheld information for over 40 years to the detriment of her long-term emotional health. Practitioners must understand that when a series of unexplainable peculiar behavioral manifestations emerge, hypotheses regarding sexual abuse should be investigated.

Within the field of gerontology, Life Review and Reminiscent therapies are held in high esteem with little critical analysis. Ardith's work provides the evidence that demonstrates the importance of in-depth social histories, which are the prerequisite for providing Life Review and Reminiscent intervention. For example, a nursing home social worker could

easily extract repressed sexual abuse memories in a manner that could induce a psychotic reaction. It is unlikely that any full time nursing home employee would have the skills or training to respond appropriately or even recognize the episode for what it is.

It does not take a college education to recognize a psychiatric disorder. However, training and insight is the prerequisite to identify causal linkages and the proper intervention–particularly if the disorder emerged as a consequence of repressed sexual exploitation. In addition, without trusting and knowledgeable practitioners, clients will continue to suffer. Here, we state *nothing new*. However, Ardith's book articulates these well accepted premises in a manner that has been absent in the literature. She provides a new vision for the treatment of sexually exploited children and adults that will significantly improve our intervention.

We feel that Ardith Trudzik's *If You Tell … It Will Kill Your Mother* is a profoundly important book for social work students, practitioners, and professors. The book provides insights that other works in this specialty are unable to achieve. We believe that this book could be required reading for anyone involved in child protective services. In addition, social work libraries should order this book. The book is very readable and as a consequence, it would fit well into the collection of most public libraries.

Barsky, A.E. (2010), Ethics and Values in Social Work: An Integrated Approach for a Comprehensive Curriculum. NY: Oxford University Press.(www.oup.com)

Reviewed by Charles Garvin, Ph.D., ACSW, Michigan Board of Social Work: Clinical and Macro License

Dr Barsky has taught at a university level since 1989 at four schools: the University of Toronto, Ryerson University, University of Calgary, and Florida Atlantic University where he is currently a full professor. He serves on the ethics committee of the National Association of Social Workers. He has written prior books on conflict resolution, interdisciplinary practice with diverse populations, as well as on social work education and on addictions. In his teaching, he has taught courses on ethics as well as social work and the law. This background has well qualified him for a book as extensive and authoritative as this one while utilizing an easy-to-read writing style and a wealth of examples embedded in well conceived classroom exercises.

This book presents a comprehensive treatment of virtually all the types of ethical issues of concern to the social work practitioner and also indicates the relevance of each of the sections to the 1999 Code of Ethics adopted by the National Association of Social Workers (NASW). The book, however, while anchored in U.S. experience, recognizes the cultural limitations of a document from one country and the author enjoins the reader to consult with other codes when it is appropriate to do so. . He also incorporates references to many other documents such as the NASW's statements about social work in end-of-life care, and about standards for substance abuse, health, school social work and other fields.

There are many strengths to this book that are often absent in other similar texts. Barsky discusses not only the ethical issues confronting practitioners working with individuals but practitioners working with families, groups, organizations, and communities or who are seeking to influence or implement social policy. He illustrates all of his major points with practice vignettes that portray the dilemmas and complexities of practice and that are representative of situations that are likely occur under contemporary political, social, and economic conditions. Every chapter concludes with similarly constructed "cases" as well as discussion questions for use in social work courses.

Barsky's use of practice examples go well beyond the practitioner's decision as to what is an ethical decision in order to show how that decision is carried out through the practitioner's interactions with the users of the social work service. Thus, the book is replete with practice examples that venture into the realm of types of interventions and processes. Here the readers who use models of practice that differ from the illustration may take different steps; nevertheless, Barsky demonstrates what ethical practice looks like and implies that other practices should operate with the same ethical constraints, even if different means to reach the same ethical ends are employed. Barsky presents a comprehensive model of ethical practice that includes an evaluation element to determine if ethical processes occurred and whether ethical ends were achieved. All of this makes the book rich in example and, thus, very readable and teachable.

The book begins with chapters that define the concepts of ethics and values and their relationships to law and the process of values clarification. These chapters also place social work values in both their historic and current context. Values and ethics are related to theories of justice as well as the theories employed by other helping professions such as medicine. Biological and psychological factors are analyzed such as moral development and some theories of moral development such as Kohlberg's are critiqued. Since spirituality is an important component of the thinking of many people, ethics in relationship to religious and

other spiritual beliefs is examined. As is true throughout the book, the use of research is integrated into this discussion and the role of institutional review boards to examine the ethics of research protocols is dealt with extensively. Barsky deals with the means of obtaining informed consent for research as well as practice procedures that will be helpful to all practitioners.

The following six chapters present value and ethical issues that arise in all contexts of practice, namely individual, family, group, organizational, community, and policy. The individual chapter considers such topics as confidentiality, competence (including cultural competence), self determination, informed consent, record keeping, and the maintenance of appropriate boundaries. The family chapter adds to this discussion the definition of who is the "client" and ethical ways of interacting with individual family members as well as the family as a whole.

The group chapter introduces the kinds of ethical issues that arise in groups and the ethical guidelines available for group work. Barsky is correct in stating that group workers tend to draw upon codes of ethics of associations outside of social work such as the American Group Psychotherapy Association and the Association for Specialists in Group Work. He is not entirely correct in referring to a code of ethics of the Association for the Advancement of Social Work with Groups. (I am the chair of the Practice Committee of that Association and we have published a set of standards for practice but have not yet created a code of ethics.) Barsky also has a well documented section on involuntary clients inasmuch as so many groups, such as those for batterers, consist of mandated clients. Many school based groups also have similar social control functions. His discussion of methods of facilitating discussion, while a good one, is another example of how Barsky illustrates ethical issues by venturing into a presentation of practice approaches (e.g., debate, dialogue).

Barsky's chapter on organizations pays attention to many interpersonal issues such as conflicts and boundaries among colleagues. The chapter has some comments on the nature of an ethical organization, as such, yet this material deserves more attention. Nevertheless, such organizational topics as record keeping, billing, and labor disputes are discussed in terms of the ethical issues they present. Similarly, the chapter on communities incorporates many appropriate topics with respect to worker actions, but I would have liked

to have seen a more general discussion as to whether or how one can speak of an "ethical community" as such. In the final chapter of this section on policy, the reader will find a full presentation of what an ethical policy is as well as the philosophies that inform ethical policy making such as egalitarianism and utilitarianism.

The last section of the book consists of chapters that Barsky terms "advanced values and ethics." This material will be especially useful to experienced practitioners who have a sound grounding in basic ethical concepts. Of use, however, to all readers is the author's "Framework for Managing Ethical Issues" as this framework identifies an entire sequence of tasks from identifying the ethical issue to evaluation and follow up after ethical decisions and the implementation of these decisions have been accomplished. Following the presentation of this model, the book offers chapters on ethical supervision and administration. The book concludes with an analysis of the ethical issues that arise in the mental health, child welfare, criminal justice, and gerontology fields. These chapters are especially useful in their discussion of concerns about client competency that arise in each of these domains as well as basic civil rights of people who are confined in various types of institutional settings such as those for children, offenders, and the elderly.

In summary, the book is very sound, comprehensive, and scholarly in its treatment of this topic. Its particular strengths lie in its linkage to major philosophical writings and typologies as well as the NASW Code of Ethics, its richness of good case examples and exercises based on these, the author's fine writing style, and the book's overall comprehensiveness. It incorporates the contemporary ideas in social work with respect to social justice and diversity. I cannot present any major criticism of the book beyond, as noted above, some places where I would have liked to see the discussion expanded although this would have been difficult to do without adding to the length of the book which is already a full-sized one.

I believe the book can and will be used in courses on ethics and values as well as most other social work courses as the latter can utilize selected chapters related to course content. I have had many years of social work practice, teaching, and research and yet found many discussions in the book that added new insights to my understanding of ethical and value issues.