PROFESSIONAL PREPARATION AND PRACTICES OF CHILD CUSTODY EVALUATORS: A COMPARATIVE STUDY OF MASTERS-LEVEL AND DOCTORAL-LEVEL PRACTITIONERS

by

AARON D. ROBB

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2013

Acknowledgements

No one completes a process like this alone, and I owe a great deal to the many people who have helped and supported me along the way. To the members of my dissertation committee from the School of Social Work, Maria Scannapieco, Randall Basham, Peter Lehmann, and Vijayan Pillai, I appreciate your encouragement and feedback during this process. A special thanks to my Committee Chair, Maria Scannapieco, for her guidance throughout this longer-than-expected journey. To David Martindale, outside committee member and long time mentor, thank you for your willingness to share your deep reserves of knowledge and the investment of your time through the final stage of this process.

To my friends and family, thank you for your encouragement and your good natured ribbing as this process continued at a sometimes glacial pace. For my parents, teachers by vocation, thank you for the lessons that put me on this track. To my wife, LeeAnn, your belief in me keeps me afloat despite whatever storms life might throw at us. To my children, if you are ever inquisitive enough to read this one day then I probably got a few things right along the way. Thank you for putting up with all those times over the years when I disappeared into my office to write – I would have much rather been playing with you too. For Melissa, my often underappreciated former assistant and proofreading guru, thank you for all the help. My mistakes are all mine, but my successes are never without support.

To my peers and colleagues, especially those in the North Texas Families in Transition Professional Workgroup, your encouragement and support have been a reminder that we are all in this together. The work we do is hard, and I am certain I might despair were it not for your reassurances. Especially to Brad, the insights you're offered over the years are a treasure beyond counting.

To the children and families it has been my privilege to serve, thank you for all you have taught me about the complexities of the human condition. I understand more keenly every day the need for humility and excellence when we are asked to intervene, and how critical it is that helping professionals understand the complexity of turbulent family restructuring.

Finally for Carolyn Kern, the first of my many professional mentors and the last of my classroom professors to issue me a grade before I completed my dissertation, I learned of your passing this morning. I am better for having known you, thank you for sharing this journey with me. You will be dearly missed by many.

November 4, 2013

Abstract

PROFESSIONAL PREPARATION AND PRACTICES OF CHILD CUSTODY EVALUATORS: A COMPARATIE STUDY OF MASTERS-LEVEL AND DOCTORAL-LEVEL PRACTITIONERS

Aaron D. Robb, Ph.D.

The University of Texas at Arlington, 2013

Supervising Professor: Maria Scannapieco

While a substantial amount of the research on child custody evaluation practices has focused on psychologists and psychiatrists, virtually none of the existing research has addressed child custody evaluation practices of social workers or other mental health professionals. This study explores and describes the training and practice differences between doctoral-level child custody evaluators, primarily psychologists, and their masters-level counterparts, who are drawn from a much wider professional background. The exact population of evaluators in the community is unknown, thus a snowball sampling method was utilized for outreach to evaluators who then provided responses to a survey instrument. Statistical comparisons of current practices across a wide variety of areas are examined. Multiple implications for practice, current policy, and future research are presented.

Table of Contents

List of Illustrations

List of Tables

Chapter 1

Introduction to the Study

Problem Statement

Divorce litigation, and increasingly litigation between never-married parents, is a reality western society has slowly been coming to deal with. One in three first marriages ends within 10 years (Bramlett & Mosher, 2001) and 43% of first marriages end within 15 years, either through separation or formal divorce (Lugaila, 1998). Roughly six out of ten divorces involve children (Clark, 1995) and, of divorces where there are children, approximately half happen within the first seven years of the marriage (Bramlett & Mosher, 2001). For never-married parents who are cohabiting, the numbers regarding separation are even higher (Bramlett & Mosher, 2001). Although there are no clear studies documenting the number of custody disputes litigating at a given time, one study from the mid 1990s noted that domestic relations disputes made up the largest type of filings, fully a fourth of the cases reviewed (Ostrom & Kauder, 1996). Hofstra Law School professor Andrew Schepard, Director of the Center for Children, Families and the Law, notes evidence indicates that the majority of domestic relations disputes are child custody related (Schepard, 2004).

Despite what modern entertainment would have the public believe, litigation is a slow and cumbersome process which is not well designed to serve the interests of parents and children. Often judges will assign a mental health professional to assist them in assessing what is in the best interests of children who are in question before the court due to their parents' legal conflict (Remley & Miranti, 1991). Sometimes this assistance is rendered in the form of clinical or

psycho-social evaluations of one or more parties to assess their mental health status. When a broader assessment regarding what parenting arrangements would be most beneficial to the child or children involved is desired (especially in the case of two adults who are well-functioning outside of their litigation issues) the family may be ordered to participate in a process called a custody evaluation (also known variously as a social study, family assessment, custody assessment, etc., depending on jurisdiction). While best available data indicates roughly 90% of cases settle without an evaluation, another 5% settle only after evaluation, with the remainder actually going to trial (Maccoby and Mnookin, 1992).

If one applies the preceding statistics on divorce and litigation to the 2.1 million people who marry every year in the United States (Centers for Disease Control and Prevention, 2013) this issue can clearly be seen as important to the field of social work due to the sheer numbers of families affected by the family court system. This is a growing population that social work professionals are called upon more and more often to provide services to (Bala, 1990). Additionally there are pervasive questions regarding the training and preparation that custody evaluators have. Many graduate programs have some training component in child abuse and neglect; however it is unclear how many of these programs deal with the much more common and equally nuanced issues encountered in suits affecting parent child relationships. There is a growing body of empirical information regarding the practices of professionals conducting child custody evaluations (Bow, 2006). While such research has primarily examined the practices of doctoral-level psychologists, it serves as a starting point for investigation of what is currently known about child custody evaluators and as a possible framework for future investigations.

Background of the Study

A review of the literature showed few clear studies of information that judges and attorneys view as important in making child custody decisions. Despite the dearth of studies regarding the practices of masters-level evaluators, one survey of family law judges (Bow & Quinnell, 2004) noted that over 40% of the time the courts appointed masters-prepared child custody evaluators. While Bow & Quinnell (2004) limited their research to one northern state, anecdotal evidence indicates that their percentages may have been lower than actual practice nationally. In one large southern state a base of county-supported Family Court Services offices exists in the most populous counties. These offices conduct the majority of child custody evaluations in the counties they serve, and according to various directors of these offices the vast majority (over 90%) of their staff do not have doctoral-level degrees (Janet Denton, Tarrant County Family Court Services Director, Denise Frank, Dallas County Family Court Services Director & David Simpson, Harris County Domestic Relations Office Director, personal communication, July 2012). Similar arrangements exist in other states. Additionally, in counties in this southern state without Family Court Services staff the courts and Domestic Relations Offices often keep listings of private evaluators. Two such publically available lists show a large percentage of masters-level evaluators amongst their lists, at 87% and 91% respectively (Bexar County Domestic Relations Office, 2013; Denton County District Court Administration, 2013).

Purpose of Study

The purpose of this study is to examine the professional preparation (i.e. graduate courses, internships, seminars) and practices (i.e., employment capacity; use of techniques such as

observation, interviews and records review; and factors considered in making recommendations regarding child custody issues) of masters-level child custody evaluators. The aim of the study is to explore how masters-level child custody evaluators practice and if that practice differs from either current practice of doctoral-level evaluators or the previously studied populations of child custody evaluators where investigative samples have primarily been drawn from doctoral-level practitioners.

Rationale of the Study

This is an important area of study for multiple reasons. The evaluative procedures employed by masters-prepared mental health professionals conducting child custody evaluations are almost completely unstudied, however both masters-level and doctoral-level evaluators draw from the same empirical literature base and ethical guidelines. While a substantial amount of the research on child custody evaluation practices has focused on the professional practices of psychologists and psychiatrists, virtually none of the existing research has addressed this issue with social workers or other mental health professionals. Indeed, several of the existing research studies have specifically excluded masters-level practitioners with an *a priori* determination that they were not relevant or of interest. Anecdotally, masters-prepared practitioners may actually provide the majority of child custody evaluation services in the country; however they may also practice in government or agency settings with a less affluent strata of the socio-economic spectrum than the practitioners who have been studied to date. If there exists an economic impact on the level of services that family court litigants receive it may be of great interest to explicate the differences. Likewise, if masters-prepared child custody evaluators are providing the same

services as their better-studied doctoral-level colleagues, but at a more affordable rate, lessons may be drawn which will aid an overburdened court system in providing more cost-effective interventions. Finally, if there is a gap in the practices of child custody evaluators based on their level of education it may be possible to examine this difference in terms of current practice guidelines and explore the potential impact of the court assigning a family to one provider versus another.

Professional Significance of the Study

The study contributes to the research literature on child custody evaluation by further examining professional practices (i.e., employment capacity; use of techniques such as observation, interviews and records review; and factors considered in making recommendations regarding child custody issues) of both masters-level and doctoral-level child custody evaluators, including comparing and contrasting between the different groups of evaluators. The study also explores demographic variables between the two groups (i.e., age, gender, years in practice), types of specialized training in child custody evaluations (i.e. graduate courses, internships, seminars), and costs associated with child custody evaluations. The data may provide better insight into the allocation of limited resources of families and courts, concurrent with the principles of social welfare and social justice espoused by the profession (Beverly & McSweeny, 1987). In over 25 years of research involving child custody evaluators this is be the first study to systematically and intentionally compare the full range of child custody evaluators, regardless of professional affiliation or degree, and reverses an ongoing narrowing of the empirical research which has moved toward focusing on only a subset of one professional group.

Chapter 2

Literature Review

Strategy for Literature Review

For the current literature review a search of multiple databases available both publicly and through the University of Texas at Arlington Library was used to locate relevant information on the subject. Library databases utilized for journal articles include Academic Search Premier, Academic Search Complete, Ingenta, Kluwer Online, PsycArticles, Psychology: A Sage Full-Text Collection, Psychology and Behavioral Sciences Collection, PsycInfo, Professional Development Collection, Social Sciences Citation Index, Social Services Abstracts, Social Work Abstracts, and SpringerLink. Keywords used for the search included "child custody evaluation," "family court," "divorce," "forensic evaluation," "parenting time evaluation," "social study," along with the individual words used in making up those phrases (i.e. "custody" and "evaluation" searched separately). Additionally, direct subscription to specialty journals (*Family Court Review* and the *Journal of Child Custody* in particular) was combined with physical searches of library archives to locate empirical studies referenced by other works.

Literature Review Results

A search of the literature yielded a large number of original empirical articles on various issues related to child custody evaluation. While many papers claimed to be empirical studies, it appeared that the vast majority of published work in this field is based on personal experience or is simply a summation of other researchers' works. Actual empirical studies fell into several

thematic categories which are discussed below. Existing empirical research on child custody evaluation appears to cover several distinct categories: simple overviews of general procedures including psychometric testing and interview types and formats; examination of bias and neutrality of evaluators; expectations of courts and legal practitioners; and specialty issues in assessment of domestic violence and sexual abuse. These issues are primarily examined through survey research, although a limited number of studies reviewed actual evaluator reports to the court. The later methodology avoided potential confounding issues such as evaluators who report behavior in line with generally accepted best practice but whose actual performance may not meet their stated standard. Direct work product records review appears significantly more limited than survey research due to the additional difficulties such reviews pose. Finally, a clear emphasis on psychologists, mainly at the doctoral-level, as the subjects of research regarding procedural issues emerged, whereas social workers emerged as the subjects of research on bias more prominently than any other group. It is unclear if the former is due to the nature of the persons conducting such research (by and large themselves doctoral-level psychologists), the convenience and ease of contacting psychologists, or other factors. The later outcome appears to be a clear artifact of the researchers and research locations for the majority of these studies, although this is not true in all cases. Please see Appendix A, Table 1, for an abbreviated breakdown comparing the groups who have been studied to this point and the purposes and methodology of those various studies.

Descriptive Survey Research

By far the majority of information regarding actual research on child custody evaluation falls into this category. The earliest work in this area begins in the early 1980s with basic and (by modern understandings) somewhat flawed articles discussing the practice of forensic social work and clinical interventions in custody and access disputes (Hughes & O'Neal 1983, Jaffe & Cameron, 1984). Hughes and O'Neal (1983) note that they failed to account for diversified practice areas (focusing instead on inpatient treatment centers) and Jaffe & Cameron (1984) note that their research was conducted in a time where the "hired gun" approach to forensic interventions was beginning to fade from common practice, resulting in mixed application of their criteria and categories. Both of these articles indicate a growing interest in understanding what was being done, in terms of common practices, in order to begin understanding best practices.

The first of what are now seen as the "modern" empirical studies of child custody evaluation begin in the mid 1980s and continue with various researchers replicating and expanding upon the earlier work (Ackerman & Ackerman, 1997; Bow & Quinnell, 2001; Gorley & Stoberg, 2000; Hagen & Castagna, 2001; Keilin & Bloom, 1986; LaFortune & Carpenter, 1998; Quinnell & Bow, 2001[1]). This set of surveys appears to form the core of ongoing research into evaluator practices. It is referenced repeatedly in studies focusing on other thematic areas such as bias and specialty practice as well as in multiple theory and opinion pieces. These surveys break down evaluator practice into areas such as interviews with each parent, the

[1] The Quinnell & Bow (2001) study is based on the same survey and sample as the Bow & Quinnell (2001) study, thus the demographic information is identical. The researchers shrewdly separated psychometric testing from other practices and were able to garner two publications from the same dataset.

children in question, the use of psychometric testing with parents and/or children, observational sessions, school and home visits, and other areas. More recently Ackerman & Brey Pritzel (2011) have conducted an update to earlier studies, with particular focus on the surveys from Ackerman & Ackerman (1997) and Keilin & Bloom (1986). Unfortunately, these most recent data have received sharp criticism over discrepancies between what was claimed in the peer-reviewed article and what was stated in the unpublished dissertation (Brey, 2007) that the article is based on. These criticisms include pointing out logical gaps in the author's assertions regarding characteristics of respondents, identification of conflicting factual statements about the nature of the responses received and their completeness, and noting the clear confusion of historical data ("have you ever…" questions) with current practices (Martindale, Tippins, Ben-Porath, Wittman and Austin, 2012). Given the extensive experience of the average practitioner who responded to the Ackerman & Brey Pritzl (2011) survey (almost a quarter century of experience on average), this last confound is one which creates significant difficulty in interpreting their data as ethical and legal requirements have evolved over the past quarter century significantly. Summaries of the data reported in these studies can be found in Appendix A, Table 2.

Other evaluator demographic factors, such as years in practice, numbers of evaluations conducted and training experiences were also examined in this series of studies (see Appendix A, Table 3) along with evaluator ratings of different factors and issues that they considered important in making recommendations to the court regarding the cases at hand. Unfortunately, this later data is not presented in directly comparable ways across studies. Keilin & Bloom (1986) developed idiosyncratic decision making factors for evaluators to rate in terms of

importance for single-parent versus joint custody decision making. These factors are not noted to be related to either any empirically derived concepts or any legal decision making criteria that the courts have to assess when rendering verdicts. Additionally, the very concept of what "joint custody" is varies significantly from jurisdiction to jurisdiction, as family law remains a state-by-state issue, making this type of rating operationally imprecise. Ackerman and Ackerman (1997) replicated this flawed rating system, drawing the conclusion

> It is apparent from reviewing these data that psychologists[2] are more careful in their decision-making process in 1996 than they were in 1986. In 1986 over 50% of the items prompted an endorsement of Parent A or Parent B. However, in the current study, less than 25% of the items resulted in endorsement of one parent over the other. (pp. 141-142)

While greater care may certainly be one valid conclusion, it rejects the effects of history and the changing understanding of the meaning of joint custody in various states. Ackerman & Ackerman (1997) could have just as easily concluded that changing application of the law (with an increase in parents sharing legal custody of their children) had led to fewer single-parent custody considerations, or that the definitions of single-parent custody and joint custody may not be the same, either in black-letter law or common use, than in the decade preceding their study.

LaFortune and Carpenter (1998) and Gourley and Stolberg (2000) likewise used similar approaches as Keilin and Bloom (1986) in terms of their assessment of decision-making factors.

[2] This study examined only the behavior of psychologists, as opposed to Keilin & Bloom (1986) who also examined psychiatrists (18%) social workers (1%) and other "masters level" practitioners (2%)

While this provided for consistency with earlier studies, it is again data of questionable foundation. Similarly, the work of Lowery (1985), while it predates the study most other researchers have chosen to replicate, creates a listing of factors that are somewhat similar to those in Keilin and Bloom and which share many of their flaws. Lowery does provide an interesting comparison between psychologists and social workers, with both professional groups reporting a substantial degree of consensus on the factors presented. Ultimately the primary contribution of these studies appears to be in data regarding other professional practice areas outlined on the various tables. Fitzgerald & Moltzen (2004), in the only article regarding child custody evaluation located from New Zealand researchers, similarly use the Keilin and Bloom factors. While this study offered some interesting insights into differences in legal systems in far-flung English speaking nations, it was primarily done as a replication of an earlier, unpublished study that was not accessible. As such it appeared to have little to offer in comparison to the broader discussion

Rather than repeating previous researchers' potentially flawed analyses of decision making factors, Bow & Quinnell (2001) tied their rating scale of decision making factors to the legal concept of "best interests of the child" as defined in Michigan State law. They presented a listing of factors as defined in law and asked their respondents to rate each on a Likert-like scale. While much like Keilin & Bloom (1986) or Ackerman & Ackerman (1997) there is no noted empirical support for these factors as having any correlation to child outcomes, this rating scale at least tied respondents to a real-world set of criteria that courts in one jurisdiction are required to grapple with. Given that child custody evaluations are designed in part to assist the courts in their

decision making, this approach appears to have better face validity than an arbitrarily selected set of factors.

While Bow and Quinnell (2001) do not reference the work of Jameson, Ehrenberg & Hunter (1997), the latter's work on operationalizing and rating factors in the "best interest of the child" concept may have had some indirect influence on the use of that criteria (rather than more idiosyncratic items) in their survey work. Jameson et al. (1997) cite both Keilin & Bloom (1986) and Ackerman & Ackerman (1997) in prompting their own development of a survey designed to assess what psychologists[3] view as important factors in the "best interests" criteria. Jameson et al. (1997) utilized a number of sources from both child development and family law to operationalize issues within the "best interests" concept and then obtained ratings of the various factors they identified. Unfortunately, while this is interesting information from a descriptive level and they provide interesting hypothetical constructs from which to make future evaluation decisions, little has been done to further support this work.

One study that appears seldom mentioned in the dialog in literature on the early assessment of issues related to child custody evaluation practices is Caplan & Wilson (1990). This study, published in a Canadian law journal, may have suffered from lack of exposure to the broader mental health community; however it appears to have taken an inclusive look at social workers, psychologists, and psychiatrists. Differences in formatting of questions and responses have led this author to include this study's results separately (see Appendix A, Table 4) but it none the less provides interesting information on both demographic and practice differences from the other samples. Most notably in terms of inter-professional differences the study noted

> Most of the respondents (82 per cent) felt that their professional discipline is relevant to their assessment approach…social workers' reasons related to their focus on environmental, relational and social factors; and psychologists' reasons were primarily related to their ability to conduct standardized testing. (p125)

Caplan & Wilson (1990) also detail evaluators who are woefully ignorant and unprepared for their tasks, sometimes willfully so (for instance, almost a quarter of respondents reporting it is unimportant to have understanding of crucial legal issues related to child custody). It is unclear if this is some bizarre artifact of Canadian practice or an issue of history (as this research occurred before most commonly accepted professional guidelines were implemented) but it clearly demonstrates any future research would be well served by explorations of basic underlying assumptions, such as understanding of the legal system is important when offering psycho-legal opinions.

Psychometric Testing

While the core studies (Ackerman & Ackerman 1997; Bow & Quinnell 2001; Gorley & Stoberg, 2000; Hagen & Castagna 2001; Keilin & Bloom, 1986; LaFortune & Carpenter, 1998; Quinnell & Bow, 2001) demonstrate a trend as to how to report procedural data, the representation of frequency of use data with psychometric testing (such as percentage of time it is used) has been challenged as over emphasizing this facet of custody evaluation:

[3] Again this survey was only of psychologists.

> … another look at the data reveals that, despite the general impression to the contrary, there is very little in the use of conventional psychological tests that rises even to the level of usual and customary, much less to the level of a standard of practice." Hagen & Castagna (2001), p. 270.

Please see Appendix A, Table 5, for examples of how calculation and presentation of usage rates based on percentage of total evaluations they are used in provides a different sense of actual frequency of use.

In regards to psychometric testing, reports indicate that, on the whole, there is a general lack of consistency between many of the clinical tests employed and the population they were developed for and normed with. Although the Minnesota Multiphasic Personality Inventory-2 (MMPI-2) and the Millon Clinical Multiaxial Inventory-II/III (MCMI-II/III) now have a research base of normative data for child custody litigants, (Flens, 2004; McCann, Flens, Campagna, Collman, Lazzarom & Connor, 2001) context-specific normative data appears to be lacking for other tests. There is also expressed concern that other inappropriate tests were being used, such as the Multiphasic Sex Inventory with litigants who are only alleged perpetrators (not admitted perpetrators as the instrument is designed for), or the Sexual Abuse Legitimacy Scale and the Bricklin Perceptual Scales, which lack any kind of empirical support at all (Bow, Quinnell, Zaroff, & Assemany 2002; LaFortune & Carpenter, 1998). The issue of inappropriate testing is compounded by difficulties in the use of those tests that could be seen as appropriate to the custody evaluation context, including over-reliance on computer-generated interpretations, lack of knowledge about interpretative cut-off points, and a failure to actually use the normative data regarding custody litigants (Bow, Flens, Gould & Greenhut, 2005). The testing issue is of

significant concern given that anyone using one of the named tests should be well aware of which populations (if any) they are appropriate (valid, normed, and adequately tested) for prior to utilizing them, and should be adequately trained in their administration and scoring. The fact that they are used inappropriately or haphazardly in custody evaluation raises concerns over the basic training of these evaluators, without even considering the more advanced issues in custody cases.

Bias/Neutrality Issues

Many writers have addressed conceptual issues of both personal bias (racism, sexism, etc.) and clinical/scientific bias (anchoring, primacy, and other heuristic phenomenon) as they impact human services delivery (Robb, 2006). The five studies identified that primarily and directly addressed evaluator bias in a child custody context, Abrams (1998), Austin, Jaffe & Friedman (1994), Caplan & Wilson (1990), Cohen & Shnit (2001), and Sagi & Dvir (1993), examined a wide range of issues, from evaluator gender to evaluator personal experience with divorce, abuse, or neglect to language and culture of origin. The studies themselves are wide-ranging in terms of the populations examined and the particular areas of interest, leading to little inter-study comparability; however each provides insightful starting points for further research.

Caplan and Wilson (1990) appear well ahead of their contemporaries, as they actually investigated not only evaluator practices (as noted above), but also personal and professional bias issues that may effect an evaluator. Evaluator beliefs regarding their own personal factors (gender, whether the evaluator had children, the evaluator's family of origin, etc.) and the evaluator's beliefs regarding the influence of one-sided communication with attorneys, social

involvement with attorneys, and other areas they labeled as "conflicts of interest" were examined through the same survey instrument used to detail their practices and procedures. See Appendix A, Table 4 for a detailed breakdown of their findings. Some of their questions are compound, unnecessarily complex, or use ill-defined terms,[4] leading to lack of clarity at times, but they are some of the first researchers to point out that many evaluators seem to deny their own humanity. While they do not cite research in detail, they accurately point out that it is necessary for mental health professionals to be aware of how their own backgrounds shape their decision making (Robb, 2006).

The obvious flaw in this type of research is, of course, that what a person reports on a survey and their actual behaviors (due to both conscious and non-conscious factors) may be discrepant. Austin, Jaffe & Friedman (1994) address this by expanding upon Caplan and Wilson's work and investigating actual evaluator decision making through analogue case vignettes. Austin et al. (1994) also expanded the survey to juxtapose issues of evaluator experience (number of assessments conducted, a factor that was not related to other demographic variables) with the basic background characteristics Caplan & Wilson (1990) examined. As could be expected, results indicated that evaluators with more experience were more likely to be aware of (and try to compensate for) bias in their evaluations. Results from Austin et al. (1994)

[4] Such as questions that children should "live with" their mothers – as if such a question is dichotomous. If asking if the child should either always or never live with the mother this might be appropriate phrasing, however except in the most severe situations children will generally live with both parents at different times.

also indicated that, in general, there were few overall bias effects from evaluator background issues in the recommendations that evaluators made in the various scenarios.[5]

The results from Austin et al. (1994) appear to be at odds with earlier research by Abrams (1988) which notes a clear distinction in responding to vignettes from clinicians who personally experienced marital disruption and those who had not. To be clear, the Abrams study sampled a small geographic selection of clinical psychologists, whereas Austin et al. (1994) (and the predecessor Caplan & Wilson (1990) study) examined a broader range of psychiatrists, psychologists and social workers and specifically focused on levels of experience in child custody evaluation. The exact scenarios offered also differed as did the general focus. Abrams appears to have been much more interested in the effects of labeling on clinical judgment (a child described as coming from an "intact" marriage, a divorce, or a high conflict marriage) than on clinician background experience, although she unexpectedly found that there were none of the expected labeling effects in her study.[6] Approaching examination of bias in this way provides an interesting insight that may be worth replicating in some manner in future studies to evaluate whether this was an effect of the sample in this study or whether taking the focus off of acting in an evaluator capacity (where there is an expectation of a great deal of scrutiny by the court and attorneys) might reveal more non-conscious bias effects of personal experience.

A unique example of a study of evaluator bias is Cohen & Shnit's (2001) examination of actual recommendations of Israeli social workers through review of court documents. Due to

[5] Two noted exceptions were a maternal decision making preference in evaluators aged 56-70 years old when compared to evaluators aged 44-48 years old, and in "move away" scenarios based on whether evaluators had, themselves, been personally abused as children or not.

statutory aspects of Israeli law there is little flexibility regarding primary placement of a young child following divorce,[7] so Cohen & Shnit (2001) chose to examine the specific family factors that impacted the amount of access fathers were permitted. Surprisingly, only one factor was identified related to the parents' actual relationship with the children (direct maternal care-giving) while several other factors that would seem irrelevant to the child's needs and/or their best interests (such as who initiated the divorce, and whether the social worker met with the father) were clearly identified as effecting the amount of access the father would have. Cohen & Shnit (2001) conclude, in part,

> The social workers seem to have recommended more extensive access to men who could ask for it, on one hand, and to whom they could relate, on the other. Both of these 'pre-conditions' were met by the better educated, higher income men who came to the interviews and who were more similar to them – middle class, academic women – in culture and socio-economic background than those who did not. (p 324)

While this study tells us little about the specific demographics of the evaluators themselves the fact that a statistical analysis of their reports yields such surprising information is instructive for future research and reflects the value of direct analysis of work product in addition to surveys and responses to vignettes.

---

[6] It is also worth noting that the vignettes in Abrams' study also elicited no differences in treatment recommendations as "[i]n general clinicians uniformly tended to recommend treatment irrespective of family condition" (p. 202). This is hardly surprising given her sampling pool.

[7] Israel operates under a "tender years" doctrine that women should be the primary caregivers of young children.

Finally, in regards to direct bias survey research one other early vignette-based study in particular pointed out a clear maternal gender bias in a specific population (female Israeli social workers) which persisted despite the fact that they are theoretically driven in the scenarios to decision making based a “best interests of the child” standard which is highly similar to that in the United States (Sagi & Dvir, 1993). This study is also notable for the antiquated and thoroughly out of date references that are used by its authors, who go so far as to call a study from twenty years prior to their publication date “recent.” Their most current reference is from ten years prior to their date of publication, and they completely ignore critical research from the early and mid 1980s. While their study is one of the earliest empirical assessments of gender bias in custody evaluation, it would have been significantly enhanced by review of its findings in light of Keilin & Bloom (1986), along with more realistic language regarding the time frames of other events to which they refer.

Although the above cited studies are the only located empirical examples where actual personal and/or clinical bias issues were examined, a common issue addressed in the research is concern over the role of the evaluator as a neutral party, rather than as an advocate for a particular side or position (Austin, Jaffe, & Friedman, 1994; Bow, Quinnell, Zaroff, & Assemany, 2002; Gourley & Stolberg, 2000; Sagi & Dvir, 1993). Even in early studies of this issue neutrality was clearly a goal of evaluators (Jeffe & Cameron, 1984; Keilin & Bloom, 1986), as would be expected given the long history of research into clinical bias (Robb, 2006). Unfortunately, as noted in the following review of actual work product, this may be another issue that exists mainly on a theoretical level, and that custody evaluators may report higher levels of attainment of professional standards than are achieved in day to day practice (Hovath, Logan, &

Walker, 2002). Even the most positive findings suggest the need for better training of custody evaluators and more standardization of the process in order to minimize clinician bias (Bow & Quinnell, 2001; Gourley & Stolberg, 2000).

Interpersonal and Sexual Violence

Three studies in the early years of this century, Bow & Boxer (2003); Bow, Quinnell, Zaroff & Assemany (2002); and Logan, Walker, Jordan & Horvath (2002); examined specialty issues of evaluator training in the areas of domestic violence and sexual abuse. This small number of studies on these types of interpersonal violence is somewhat surprising, given the prevalence and profound impacts of violence in relationships, especially in terms of risks to children (Appel & Holden, 1998; Ayoub, Deutsch & Maraganore, 1999). This may be in some part due to the ongoing political conflicts between various research paradigms regarding defining family violence and the resultant difficulties in operationalizing some of these questions (Bow & Boxer, 2003; Ver Steegh & Dalton, 2008). As of the turn of the century the overwhelming majority of states had some requirement that courts (and thus child custody evaluators) consider issues of domestic violence in making determinations, including presumptions limiting legal rights of access to children for perpetrators of violence (Lemon, 2000).

The results from the various studies looking at violence are mixed. While Bow & Boxer (2003) note adequate training and multi-source data collection in regards to domestic violence, this exists in stark contrast to Logan et al. (2002) who, in reviewing actual evaluations,[8] note a

[8] More on the differences between reported and actual behaviors is discussed below in reviewing work product analysis as a whole.

> lack of attention to domestic violence [that] raises serious questions about evaluators' understanding of the risk of harm for children and parents in cases of domestic violence. (p. 736)

It is difficult to draw definitive conclusions between the two studies, as Bow & Boxer (2003) surveyed a national sample of evaluators, while Logan et al. (2002) focused on only one sparsely populated county in Kentucky. Bow et al. (2002) were much more critical of evaluators' understanding in terms of sexual abuse, however, noting that even in their self-reports evaluators failed to describe the use of any protocols, models, or guidelines, let alone any empirically validated information in their assessments of sexual abuse cases. Indeed, as has been noted in earlier general surveys of psychometric testing where there tests without empirical support were used, or tests were used inappropriately (Quinnell & Bow, 2001; LaFortune & Carpenter, 1998) evaluators reported using psychometric tests designed for confessed or convicted offenders with alleged offenders as if the two groups were normatively equivalent, along with "tests" where the publishers have refused to release underlying data for external studies of validity or reliability. This was thankfully balanced by the evaluators' rejection of several unsupported concepts and theories regarding identification of sexually abused children, although Bow et al. (2002) note

> a relatively high percentage of respondents were unfamiliar with the concepts, which is troublesome considering they are often discussed in the sexual abuse literature. (p. 571).

Overall the actual level of training and expertise brought to bear in child custody evaluation cases involving physical and sexual violence remains an unclear and under-studied issue.

Legal System Expectations

Perhaps most interesting of all the survey research studies published were the four that focused on expectations of judges and attorneys and how that might inform evaluator practices. In an environment such as the courtroom there is clearly a multidisciplinary interaction amongst professionals, and it seems remiss to ignore how the expectations of the legal profession may impact on the mental health professionals brought into the legal area to assist in providing custody evaluations. The earliest of these reports, Crosby-Currie's 1996 survey of mental health and legal professionals is (unlike its earlier contemporaries) less an overall analysis of evaluator practices and instead, it focuses more on the differing importance that evaluators, attorneys and the judiciary place on different methodologies by which children are involved and the relationship between children's ages and their involvement. Crosby-Currie (1996) explores the case law in various jurisdictions on children's input to the court process and provides a broad survey of multiple jurisdictions in regards to actual expectations of the various professions. While Crosby-Currie's work has a clear flaw in limiting its survey of mental health professionals to only clinical psychologists (many of whom may not have been involved in child custody issues), it does manage to cover all family law judges in two states in its sampling frame with a better than 38% response rate reported. While there were expected differences between each of the professional groups, responses indicated that the likelihood a child would be asked about

their wishes and the subsequent weight given those wishes increased with the age of the child. In terms of interviews of children, both judicial and mental health professionals reported similar preferences regarding various types of questions, although some questions remained regarding whether mental health professionals and legal professionals were using similar meanings when discussing technical interview issues.

The remaining three studies, one of attorneys and judges (Bow & Quinnell, 2004), one a reanalysis/comparison of data from various surveys of attorneys, judges, and psychologists (Ackerman, Ackerman, Steffen & Kelley-Poulos, 2004), and the other of judges alone (Ackerman & Steffen, 2001), appear to have been conducted much more in line with the existing research on evaluator practices. This seems to be, in large part, due to the researchers involved and their previous publications in this area. The Ackerman, Ackerman, Steffen & Kelley-Poulos (2004) paper is a comparison of the Ackerman & Ackerman (1997) data, the Ackerman & Steffen (2001) data discussed below, and a survey of family law attorneys, purportedly by Ackerman & Kelly-Poulos (2001). Unfortunately, this last survey which makes up the paper is not listed in the paper's references, and it is inconsistently date-referenced in the paper itself, thus much of the information that might be useful here is unsourced and not utilized in this review.

Ackerman & Steffen (2001) directly replicate the earlier Ackerman questionnaire and compare various judicial responses to those of respondents in the Ackerman & Ackerman (1997) and Keilin & Bloom (1986) studies. Bow & Quinnell (2004) instead offer simple rankings of the information that judges and attorneys view as important. Both surveys indicated a strong preference for the child custody evaluator to serve as a neutral expert (either court appointed or

as mutually agreed upon by the parties involved), which appeared in line with the generally prevailing understanding of professional best practices. Each study offered detailed analysis of views from the bench about what issues were most important to the court (which showed significant but not exact overlap with the issues viewed as important by evaluators). Bow & Quinnell (2004) also noted that attorneys reported that the court generally retained masters-level evaluators (social workers and psychologists) 42% of the time when a child custody evaluation was needed, with doctoral-level psychologists (51%) and psychiatrists (7%) making up the remainder of the evaluators. This is the only statistic noted in any of the research regarding who courts select for evaluations, and it seems to reflect that the primary focus on doctoral-level psychologists as subjects of research may be a substantial mismatch with those who are actually performing evaluations.

Work Product Records Reviews

As has already been pointed out, the obvious flaw in survey research is, of course, that what a person reports on a survey and their actual behaviors (due to both conscious and non-conscious factors) may be discrepant. Eight studies were identified that specifically addressed review of actual evaluator work product: Bow & Quinnell (2002); Brandt, Dawes, Africa & Swartz (2004); Cohen & Shnit (2001); Davidson-Arad, Cohen & Wozner (2003); Horvath, Logan, & Walker (2002); Kunin, Ebbesen & Konecni (1992); Logan, Walker, Jordan & Horvath (2002); and Schindler (1985). The Cohen & Shnit (2001) and Logan, Walker, Jordan & Horvath (2002) studies have already been described above in dealing with issues of bias and violence, respectively.

The Kunin, et al.(1992) study, although using case information from ten years prior to its publication date, produces one of the most stunning pieces of research in the field: a mathematical model of judicial decision making as affected by evaluator recommendations and a breakdown of factors that influence evaluator recommendation. The research was able to demonstrate that evaluations were significant factors in judicial decision making with a clear line of causality from evaluator to judge (that is, the evaluator's recommendations influenced the judges' decisions, rather than merely anticipating correctly what the judges would have done anyway). The log-linear and factor analysis data presented in this study may be somewhat daunting to the less mathematically inclined, but the line of reasoning presented is quite clear. Unfortunately the data reviewed for this study is from a time where sole custody to one parent was the predominant recommendation, instead of the joint custody paradigm which emerged in the 1980s and has led to the more modern idea of shared parenting time, rather than treating children as objects to be possessed by one parent or the other. Additionally, due to idiosyncratic issues with the judicial district sampled, it is unclear if the type of evaluations used in 1982 are anything like the child custody evaluations seen since the development of multiple professional guidelines for this area of practice.

Two much more contemporary (and conflicting) work product reviews come from Bow & Quinnell (2002) and Horvath, Logan & Walker (2002). Both studies reviewed a relatively small number of cases (52 and 82 respectively) with Bow & Quinnell (2002) focusing solely on doctoral-level psychologists, and Horvath, Logan & Walker's (2002) sample being approximately three-fourths social workers and one-fourth doctoral level psychologists. The majority of the social workers were direct employees of an office of the court who conducted

evaluations for little or no cost to the parents involved. Each study referenced the current state of guidelines and the self-report survey work that had led them to choose to examine actual reports, noting that the self-reported congruence with guidelines may not play out in actual practice. These reports differed over whether actual compliance with guidelines could be substantiated from simple report reviews, with Horvath, Logan & Walker (2002) maintaining it was not possible to judge adherence to the APA guidelines from just the written report and instead incorporating additional factors from family law publications. Bow & Quinnell (2002), on the other hand, incorporated review of both APA and AFCC guidelines. Neither study addressed APSAC or other relevant guidelines. It appears clear that this is a point of departure for these different sets of researchers in how they operationalized their checklists and it is not surprising that they end up drawing different conclusions in their analyses. While Bow & Quinnell (2002) find relative congruence with previously reported data on evaluation practices, Horvath, Logan & Walker (2002) conclude there is a lack of consistency between guidelines and practice. This is unsurprising, as there are some acutely obvious differences in the results that they reported; for instance Bow & Quinnell (2002) report 100% of their sample had an "interview with mother" whereas as few as 78.1% of the private (non-court-employed) evaluators in Horvath, Logan & Walker (2002) "assessed mother." It is unclear if the two studies are using the words "interviewed" and "assessed" similarly, and if so how the subjective call would be made that an evaluator failed to assess a parent that they had interviewed, but it does seem a critical failure to meet professional standards for almost a quarter of private practitioners to ignore the assessment of one of the parents involved. Both of these studies beg for further replication and clarification of how certain areas were functionally rated.

In their study of South African psychologists Brandt, Dawes, Africa & Swartz (2004) chose to review reports to examine the substantive issues that inform decision making in custody evaluations. Using thematic content analysis they derived a number of themes regarding factors that custody evaluators found important in their decision making. The researchers then formulated operational definitions of these themes and ranked them according to level of occurrence in the reports that they reviewed. While this data is somewhat idiosyncratic due to the differences in practice and judicial process in South Africa (compared to other western-style jurisdictions), the approach itself is insightful and clearly highlights areas of parenting that the evaluators believed impacted the family sufficiently to incorporate it into their reports. Rather than compare any of these results to previously studied survey data the authors merely present and summarize, urging much caution in interpretations due to small sample size. This leaves their work begging for replication with a more substantial sample, but it serves as an interesting starting point for further contextual analysis.

The final two work product studies, Davidson-Arad, Cohen & Wozner (2003) and Schindler (1985) examined similar issues of the importance of various decision making factors as determined via reports to the court from evaluators in Israel. As noted already, the Israeli legal context presents substantial differences in underlying legal assumptions from jurisprudence in the United States, however there are no clear indications how this would then change the relative importance of various decision making factors within that context. The Schindler (1985) study suffers many of the same issues as other research conducted on samples from the late 1970s and early 1980s that has been remarked on previously, along with issues of small sample size, although it is groundbreaking in both its format and its purpose. The Davidson-Arad et al. (2003)

work focused more specifically on the influence of perceived quality of life issues (as assessed by Shye's Systematic Quality of Life Model) and how quality of life interacted with parental features. Schindler (1985) identified seven categories that evaluators viewed as important, however it was unclear if these were thematically derived or recurrent *a priori* categories that researchers placed various report elements into. These categories appeared to be of little predictive value however, as they appeared in less than a quarter of each of the reports, although there was some congruence with previously reported survey work. The later work of Davidson-Arad et al. (2003) provided a much greater level of detail and created a purposeful sample of reports where each evaluator submitted the most recent primary-maternal custody and primary-paternal custody recommendations. This resulted in a sample where both maternal and paternal decision making factors could be fully examined, although the 50/50 split in primary custodial parent gender did not reflect the predominantly maternal primary custody reality the researchers note as existing in Israel. Using discriminate analysis and basic chi-square tests for categorical variables the researchers explored the significance of the various quality of life variables versus various parental characteristics, finding that predicting custody recommendations based on the quality of life issues was the most accurate method. There were also strong indications of a gender bias operating in the different recommendations, leading the researchers to posit that problems for fathers were weighed more heavily against them than the same problems for mothers due to an unspoken expectation on the part of evaluators that mothers would provide a higher quality of life simply because of gender. These fascinating findings, backed by strong statistical support, provide clear directions for further research as well.

Evaluator Demographics

In the primary studies to address both evaluator practices and demographics (Ackerman & Ackerman 1997; Ackerman & Brey Pritzl, 2011; Bow & Quinnell, 2001; Gourley & Stolberg, 2000; Keilin & Bloom 1986; LaFortune & Carpenter 1998; Quinnell & Bow 2001) a clear pattern emerges that an overwhelming number of the participants surveyed are doctoral-level psychologists (see Appendix A, Table 3). Despite the number of studies focusing on evaluator bias and specialty training in the United States where a substantial number of the participants were social workers (Austin, Jaffe & Friedman, 1994; Bow & Boxer 2003; Caplan & Wilson, 1990) and a number of studies outside of the United States specifically on social workers they, along with other masters-level practitioners, are largely ignored in general review literature.

Training and experience are also areas of minimal focus in broader studies of evaluator practices and procedures, although it appears that this facet has received greater attention in the most recent studies. Despite the existence of research indicating that general training and experience are poor predictors of competence, as referred to in LaFontaine & Carpenter (1998), it appears that many custody evaluators in their study valued formal credentials and licensure over forensic experience or involvement in research. LaFontaine & Carpenter argue that

> …<u>specific</u> training and experience may be more valuable, and those active in research are arguably most likely to be current in assessment strategies and the validity of literature (p 213). [<u>Emphasis</u> added]

This is a theme that repeats itself, although never quite as clearly or directly, throughout the reviewed research.

This idea is supported in other studies that indicated specific experience in child custody evaluation leads to more extensive evaluations. The more experienced an evaluator is the more it appears that they are likely to use additional information outside of the direct interview with the parties involved in the decision making (Austin, Jaffe, & Freidman, 1994). In addition, it appears that more complicated case issues receive more time for review of additional information (Bow, Quinnell, Zaroff, & Assemany, 2002). A note of caution is sounded regarding clinician's self assessments, however, when they rate themselves highly due to their experience but then tend to express some doubt about their colleagues' performances. It also appears that less educated direct employees of the court may adhere more strictly to professional guidelines than terminally degreed clinicians in private practice (Gourley & Stolberg, 2000; Hovath, Logan, & Walker, 2002).

Chapter 3

Framework

Definitions and Research Approach

The language regarding child custody is not uniform across the United States, and it has changed over time at different rates in different jurisdictions. As part of this research it appeared imperative that we begin with common definitions, especially where overlapping terms may hold different meanings to different subjects.

Child Custody Evaluation

Historically there have been a number of gender-based presumptions about the care of children after the dissolution of a marriage, however a push for gender-neutral approaches to legal decision making and changes in no-fault divorce laws, coupled with societal changes such as increasing involvement of women in the workplace and men in child rearing, have led to case-by-case consideration of the needs of the children in question (the “best interest of the child standard”) becoming the governing legal doctrine in legal decision making (Pruett & Barker, 2009). Often courts appoint a neutral mental health professional to provide an evaluation of children and parents, along with an assessment of the family dynamics and what arrangements would be in a child’s best interests (Remley & Miranti, 1991, Schepard, 2004). Neutrality in the role, even when operating in a jurisdiction which permits each side to select “their own” evaluator, rather than providing advocacy and litigation support to one side, separates evaluators from consultants. Perceived objectivity has been acknowledged as a critical component of

evaluations from the beginning of studies on evaluators (Keilin & Bloom, 1986). Commonly such evaluations are referred to collectively as child custody evaluations, although they have historically been referred to by names such as social studies (now seen as an archaic term, see commentary to this effect in the Uniform Child Custody and Jurisdiction Enforcement Act, 1997), home studies (often in confusion with foster care evaluations or evaluations for dependency courts, see for instance Dickerson, Allen & Pollack, 2011), and more recently parenting time evaluations (as some jurisdictions have moved towards more plain-language statures governing post-separation parenting arrangements). For the purposes of this research, evaluations by a neutral mental health professional regarding allocation of post-separation parenting arrangements were collectively referred to as child custody evaluations in order to eliminate confounds from differences in jurisdictional legal terminology, regional variations in reference frames, and historical differences in terminology.

Joint Custody

Joint custody is an imprecise term at best, as it may mean different things from jurisdiction to jurisdiction. Additionally there are clear differences between referring to joint legal custody, assessing rights and responsibilities a parent has in regards to a child, and joint physical custody, where children may spend large portions of time with each parent (Pruett & Baker, 2009). Compounding the definitional difficulties, legal joint custody can further be divided up in ways where parents may share certain rights and responsibilities at all times, have certain rights and responsibilities only when the child is in their care, or legal authority to make decisions regarding specific areas of the child's life, such as schooling or medical care, may be

differentially divided between the parents (Warshak, 1992). Schepard (2004) loosely defines joint custody as

> "a post-divorce parenting arrangement in which parents substantially share decision-making for their child (joint legal custody) and the child spends substantially equal time at each parent's residence (joint physical custody)" (pg 35).

This is an adequate broad definition of a term which may, in practice, mean different things to different people. For purposes of this research questions regarding joint custody were addressed by specifically asking about either decision making (shared or otherwise) or in regards to time under the care of a particular parent, with clear reference to Schepard's definition as quoted above. When more targeted questions were asked regarding specific caretaking arrangements or legal authority, the questions were operationalized in as behavioral a manner as possible.

Parenting Time Arrangements

Many different groups, such as the American Academy of Matrimonial Lawyers, and the Association of Family and Conciliation Courts at the national level, and the Alaska Court System and the State Bar of Arizona at the state level, have proposed model schedules for children spending time in two parents' homes (Lamb & Kelly 2009). These collections of time-sharing options are referred to under various terms such as "Model Parenting Agreement," "Model for a Parenting Plan," "Model Parenting Time Plans," and "Planning for Shared Parenting." These models reflect an ongoing shift away from a dichotomous approach to parenting time, where one parent "has custody" of the child, or is the "primary parent" and the child "visits" with the other

parent, toward a healthier approach where the child's relationships with each parent are addressed (Pruett & Baker, 2009, Schepard, 2004). While some authors have referred to parenting time arrangements as "placement schedules" (Ackerman & Brey Pritzl, 2011), this appears to be an idiosyncratic anomaly of jurisdictional terminology more reflective of placement in a foster care system than parenting a child between two homes. As such, and consistent with the broader literature, references to schedules of parenting time were made by discussing actual scheduling rather than as "placement" activities or using the outmoded "primary vs. visiting" parent terminology.

Substantially Equal Schedules

Consistent with the evolution in approaches to parenting time arrangements, additional nuance has entered the discussion regarding shared parenting arrangements where there may not be an exact 50% split in parenting time between parents, but where the difference in the amount of time spent with one parent or the other is negligible. Ackerman (2008) has noted that an increasing number of jurisdictions use the phrase "substantially equal" to describe this type of schedule. There are concerns that there are still many parents who are focused more on an exact 50% split in time, rather than a less rigid approach that may be in the child's best interests, due to the financial implications that are tied to such divisions of time (Ackerman & Brey Pritzl, 2011). This raises new questions for study, such as if evaluators see a 55/45 split of parenting time (or even larger deviations from a rigid 50% split) as a distinction without a difference. Although there have been previous anecdotal claims in this regard (Ackerman, 2008), to date only Ackerman & Brey Pritzl (2011) have addressed this issue with actual data.

Ultimate Issue

Although evaluation of what arrangements would be in a child's best interests (Remley & Miranti, 1991, Schepard, 2004) have long been seen as core issues for child custody evaluators, there is also an ongoing debate in the professional literature as to what level of detail can be reliably provided to the courts regarding this issue. Tippins & Wittmann (2005) proposed a four-level stratification of data and resulting inferences that evaluators could draw:

- Level I, what the clinician observes (i.e. basic descriptions if interactions and behavior)
- Level II, what the clinician concludes about the psychology of a parent, child, or family (i.e. abstractions regarding an authoritarian parenting style, or depression)
- Level III, implications of Level II conclusions for custody-specific variables (i.e. fit between children's behavior and parents' empathy and capacity)
- Level IV, custody-related "shoulds" (i.e. prescriptive schedules involving value judgments between competing scenarios)

At each of these levels there are various levels of empirical support, with arguments being made that at the top levels a clear evidence base to scientifically draw such conclusions is lacking (Tippins & Wittmann, 2005). Other authors have argued it is not necessarily what question is being answered (the ultimate issue before the court or a more specific issue of behavioral significance) but rather the analytical gap that may lie between the available data and the conclusion reached by the evaluator that poses the greatest problem (Zervopoulos, 2008). Existing research literature has also produced some interesting results when attempting to operationalize these type of questions, such as divergent responses over whether evaluators should answer ultimate issue questions, but agreement from these same responders regarding

stating who is the "better parent." In the end this is seen as perhaps a distinction without a difference for the court (LaFortune & Carpenter, 1998) despite whatever semantic contortions are used.

Regardless of what theoretical stance is adopted on the reliability of ultimate issue recommendations in child custody evaluations, there is a clear utilitarian argument that it is important to understand how evaluations are occurring, including what recommendation types are being offered. Part of that process for this research involved questioning evaluators regarding not just whether they are offering "ultimate issue" recommendations, but what level of recommendations and conclusions they feel appropriate to provide.

The Emerging Standard of Care

In the ongoing absence of clear empirical data regarding effective parenting behaviors (DeClue, 2002), and even the willful misuse or distortion of research for political purposes (Milchman, 2000; Warshak 1996), a number of professional guidelines for forensic experts continue to serve as the formal standard of practice. Current professional forensic guidelines are structured to promote objectivity in assessment, with the ultimate awareness that there are multiple consumers of information involved in the evaluative process. While the litigants are owed various duties of care by evaluators, they are not the "client" of the forensic expert in the same way a person receiving clinical services might be (American Professional Society on the Abuse of Children, 1990; American Psychological Association, 1994; Association of Family and Conciliation Courts, 2003). While each professional body has different areas of interest or specialization, the goal remains to provide as accurate a report as possible. Multiple guidelines

exist, from those promulgated by interdisciplinary committees from legal groups (the Association of Family and Conciliation Courts and the American Academy of Matrimonial Attorneys), to single-discipline professional associations (the American Academy of Child and Adolescent Psychiarty, the American Psychological Association, and the National Association of Social Workers, Oregon Chapter) and myriad individual state statutes and regulations. While there is clear acknowledgement in many of these resources that masters-level evaluators from multiple disciplines exist, the most recent research has focused on the practice of psychologists and adherence to the APA Guidelines (Ackerman & Brey Pritzl, 2011; Bow & Quinnell, 2001; Quinnell & Bow, 2001). While there is substantial overlap and agreement between guidelines, particularly those from interdisciplinary groups, it remains unclear if adherence to professional standards is typical outside of the previously-studied doctoral-level psychologists.

The Problem of Historical Confounds

One of the practical problems in comparing professional practices across different times is that historical changes influence behavior (Monette, Sullivan & DeJong, 2002; Rubin & Babbie 2008). There have been ongoing evolutions regarding child custody issues, both in professional understanding and societal expectations, and as such it may be difficult to draw firm conclusions in comparing current respondents to those who have participated in the previous published works without acknowledging this issue. Additionally, there appears to be a recurrent flaw in the literature regarding how questions about *current* practices have been asked. While Bow & Quinnell (2001) specifically asked respondents about the practices they utilized the "vast majority (>75%) of the time" (James Bow, personal communication 2/18/2008), and Keilin &

Bloom (1986) reported they asked what constituted typical procedures, other research has not provided that clarity. Ackerman & Brey Pritzel (2011) have been rightly criticized for presenting information obtained from questions asking "have you *ever*" as current information (Martindale, Tippins, Ben-Porath, Wittmann & Austin, 2012). Clearly an evaluator who utilized an examination procedure at the start of their career decades ago, but discarded it as the evidence base changed or they became more knowledgeable, should not be reported as "currently" using that procedure. Other researchers have had similar issues (although none have attempted to obfuscate them), reporting the frequency of use of various procedures that had "ever" been used (Ackerman & Ackerman, 1997) or the average frequency of use, without explicitly specifying if that was current use (LaFortune & Carpenter, 1998). It appears critical if one is to obtain information regarding current practices that questions specifically ask about current practices.

As a side issue, it remains unknown as to whether comparisons of responders in the current research to their historical counterparts, treated as a longitudinal trend study, might produce novel information in order to contrast responses from doctoral-level responders to existing research. Unfortunately no such longitudinal approach seems feasible with the non-doctoral level responders, given the paucity of their representation in extant research. That said, inquiries of the responders regarding if they have completed surveys for previously published research were utilized to give an indication of the uniqueness of responses received and/or the reach of the snowball effect of the survey. As there is no way of assessing whether the responders who report participation in past research were part of the actual data set used (several studies excluded participants at different experience levels, or discarded survey responses which were otherwise unusable) it is difficult to support any direct comparison, such as in a true panel study (Rubin &

Babbie, 2008). Given the historically small number of responders and the unknown size of the population of custody evaluators, care will need to be exercised in any conclusions drawn.

Chapter 4:

Methodology

Purpose

The purpose of the study is to examine the professional preparation (i.e. graduate courses, internships, seminars) and practices (i.e., employment capacity; use of techniques such as observation, interviews and records review; and factors considered in making recommendations regarding child custody issues) of child custody evaluators. The aim of the study is to explore how masters-level child custody evaluators practice and if that practice differs from current practice of doctoral-level practitioners.

Research Design

This study used a non-experimental survey approach to obtain exploratory and descriptive data. When conducting exploratory research surveys are a common methodology which allow for simultaneous analysis of multiple variables even when causality might not be known (Rubin & Babbie, 2008). A large number of cases is important in descriptive analysis, and survey research makes large samples feasible (Rubin & Babbie, 2008). The study utilized a non-probability sample with the unit of analysis being individual evaluators. In the interest of efficiency, and in order to minimize time and historical effects on subjects, a cross-sectional approach was used with data being collected for in the late summer and early fall of 2013. Survey responses themselves were collected anonymously to encourage candid answers from participants.

Research Hypotheses

Existing research indicates a trend towards greater complexity in child custody decision making, and greater time spent in most related activities (Ackerman & Ackerman 1997; Ackerman & Brey Pritzl, 2011; Bow & Quinnell, 2001; Gourley & Stolber, 2000; Keilin & Bloom, 1986; LaFortune & Carpenter, 1998). It was expected that results from this sample would be closely reflective of current practices as reported in more recent studies. One of the areas that has been under-assessed is home visits, with only two of the major studies reporting data on this procedure, whereas psychometric testing has received much attention and substantive criticism over the flawed ways in which data has been reported (Ackerman & Ackerman 1997; Ackerman & Brey Pritzl, 2011; Hagen & Castagna, 2001; Martindale, et al. 2012). It is expected that this has occurred as a function of selection bias in earlier studies due to their exclusion of the full spectrum of child custody evaluators. Consistent with these assumptions, the following hypotheses will be tested in this study:

1. There will be no significant difference between doctoral-level evaluators and masters-level evaluators regarding time spent on interviewing, parent-child observation, and records review.
2. Doctoral-level evaluators will focus more time on psychometric testing than masters-level evaluators.
3. Masters-level evaluators will focus more time on naturalistic observation (home visits) than doctoral-level evaluators.
4. Masters-level evaluators will be more likely to be employed in an agency setting or affiliated directly with a court services office than doctoral-level evaluators.

5. Costs for services by masters-level evaluators will be significantly lower than for doctoral-level evaluators.
6. There will be no significant difference between the specialized training received by either group of evaluators.
7. Both groups of evaluators will report similar ratings on various decision making variables considered in making recommendations regarding child custody issues.
8. More experienced evaluators will spend more time in assessment than less experienced evaluators, regardless of background.

Sampling

As there are no official listings or compilations of child custody evaluators a non-probability sampling procedure is necessary (Rubin & Babbie, 2008). Because of the diffuse affiliation of child custody evaluators, snowball sampling was utilized in order to reach the broadest section of the population of child custody evaluators possible. Similar to the most recent study on evaluators (Ackerman & Brey Pritzl, 2011) contacts were made to professional associations catering to child custody evaluators to obtain e-mail contact information regarding their membership. A letter explaining the purpose of the study and requesting their participation was sent along with a link to a website containing the survey (example copy attached in Appendix B).

Additionally, public agencies and associations of public agencies providing child custody evaluations (primarily Family Court Services offices attached to various courts), publicly available listservs frequented by child custody evaluators, and national professional associations

were contacted with a similar letter and link to a website containing the survey. All of these communications requested that the recipient pass the information on to others who they know provide child custody evaluation services and to re-broadcast the request for participation as widely as possible.

To be included in the sample a respondent had to meet the following inclusion criteria: have part of their current practice devoted to child custody evaluation and have access to the internet. Given that the population of child custody evaluators is an unknown, there is no way to judge the representativeness of the sample obtained. Past studies have ranged from a sample size of 82 to 213 respondents (see Appendix A, Table 3), although more recent studies have been criticized due to using an open sampling procedure (snowball sampling) and later attempting to make claims about characteristics of the population sampled from which were simply not possible to make given such an open procedure (Martindale, et al., 2012). Even if the population of evaluators was known, there is no statistical basis for judging the adequacy of response rates, although with a known population it might be able to explore response bias (Rubin & Babbie, 2008).

A minimum sample size, based on projected data analysis (see below), required at least 128 participants, with 64 participants being masters-level evaluators and 64 being doctoral-level evaluators. Beyond obtaining the minimum-necessary number of respondents for adequate statistical analysis, due to the various factors that must be taken into account, such as the exploratory nature of the research, what is feasible given the population to be studied, and time and resource constraints, success in sampling thoroughness is judged based on previous work in this area.

Protection of Research Subjects

Approval for study was obtained through the University of Texas at Arlington Institutional Review Board (IRB) prior to the beginning of any research. All data obtained was recorded anonymously and without any personally identifying information outside of broad demographic variables. Informed consent was obtained from the participant at the beginning of the survey. In order to avoid prejudicing survey responders or introducing other possible confounds, the notification indicated that we were seeking information from evaluators at all levels of practice and did not discuss the intent to compare practitioners of different educational levels. There were no appreciable concerns regarding deception through this practice as participants will be informed that this will be a comparative study, and different demographic variables will be used in the analysis. In doing this we simply sought to avoid emphasizing one demographic variable over another.

Data Collection

A hybrid of the adapted questionnaire used by Ackerman & Brey Pritzl (2011) and the Bow and Quinnell (2001) questionnaire was used in order to obtain as much comparative information as possible regarding the subjects. The questionnaire remained divided into Keilin and Bloom's (1986) original four sections, demographic information, custody evaluation practices, decision making procedures, and recommendations in evaluations. The continued modernization of the questionnaires espoused by Ackerman & Brey Pritzl (2011), addressing the subject's expectations regarding review of various records, psychometric testing, home visits, and consultations was preserved. Additionally, the approach to questions regarding psychometric

testing was clarified in order to better reflect the criticism of previous over-estimations and misrepresentations regarding psychometric testing by Hagen & Castagna (2001) and Martindale (2012). Finally, an additional question regarding if the respondent has ever participated in a survey such as this before was included in order to note when we have captured new/unique individuals. This may also perhaps permit an "apples to apples" comparison regarding the current practices of those who have ever responded to previous surveys on later data analysis. See Appendix C for the questionnaire.

Participants completed the survey on-line via website host "surveymonkey.com." The web service was set up to produce a data set which can be further analyzed via statistical software.

Data Analysis

Based on the questions examined in this study, initial analysis of many differences between doctoral-level evaluators and masters-level evaluators can best be accomplished through the use of t-tests, as the mean and variance in multiple areas are unknown (Gravetter & Wallnau, 2004). Further analysis can be attempted using ANOVA procedures to explicate differences. A significance level of $\alpha = 0.05$ is suggested as a common, conventional setting (Gravetter & Wallnau, 2004) for each of these analyses.

T-test and Effect Size

The *t* statistic is an appropriate way to test hypotheses in an independent-measures research design where the goal is to evaluate the difference between two groups (Gravetter & Wallnau, 2004). The groupings in these analyses would be doctoral-level and masters-level evaluators,

with the response variable (dependent variable) being time spent in each activity (interviewing, records review, etc.) or costs, as appropriate to the hypothesis question. For the various hypotheses comparing doctoral-level and masters-level evaluators, a two-tailed t-test will be utilized, as no *a priori* assumptions regarding the groups of evaluators are made and there is no previous research to support any such assumptions. The assumption of homogeneity of variance (necessary to evaluate for independent-measures *t*-tests) is assessed through Levene's Test for Equality of Variances.

Utilizing Cohen's *d* in conjunction with the *t*-test will not only permit conclusions regarding if there is a difference between the two studied groups, but if so the relative effect sizes of those differences. Additionally, analysis of percentage of variance explained (r^2) will be utilized using Cohen's (1988) criteria for evaluating effect size. Given the exploratory nature of this study, being able to provide more than simple binary answers regarding if there are differences in the groups may be helpful. In regards to sample size, Cohen (1992) notes "to detect a medium difference between two independent sample means…at $\alpha = .05$" (p158) requires at least 64 respondents in each group. A "medium" effect size in this case is categorized as $d = 0.50$.

ANOVA

An independent-measures ANOVA allows for multiple comparisons of various conditions between different groups. For ANOVA procedures the independent variables (technically quasi-independent variables, as they are nonmanipulated) assessed are categorical, such as educational level (the primary difference of interest discussed so far), but can also include gender, ethnicity,

or even practice settings (private practice, University clinic, public mental health agency, court-connected services, etc.). It was expected responses to the survey continuous variables, such as years of child custody evaluation practice experience, may be grouped and coded into three to four levels depending on the range of responses. Dependent variables would then be responses to decision making criteria (variables drawn from scores on Likert-like questions with responses ranging from 1 to 9, with 1 indicating the factor was not at all important and 9 indicating the factor was extremely important, treated as continuous for purposes of analysis). Additional dependent variables will be number of approaches utilized and number of hours in assessment. Similar to calculations regarding sample size for *t* tests, Cohen (1992) notes that for analysis involving two groups, at $\alpha = 0.05$ with a medium effect size ($f = 0.25$).

Chi-square Tests for Association (Independence)

Although chi-square tests for association (independence) were considered as one method of analyzing data, review of the number of subjects needed for even a simple 2x4 table (with a resulting 3 degrees of freedom) showed this was largely impracticable. Keeping the same basic approach as with previous methods ($\alpha = 0.05$ with a medium effect size) such an analysis would require 242 subjects total, half from each group (Cohen, 1992). Given that this number is almost the same as the maximum number of evaluators ever reported studied in the previous literature it during the design of the study it appeared more feasible to focus on other statistical methods. Even with a limited response rate it should still be possible to detect large effect sizes when comparing limited numbers of nominal categories.

Chapter 5:

Analysis of Results

The purpose of this study is to explore multiple factors amongst child custody evaluators of both master-level and doctoral-level education. This chapter includes a description of the study sample, comparative summaries, and research findings and results. The Statistical Program for Social Sciences (SPSS), version 21, was utilized to aid in the analysis of the data. Tables and figures are utilized to present information where applicable.

Description of Sample

A total of 305 responses of varying completion were received through the SurveyMonkey system. This data was downloaded in spreadsheet format and examined for issues of completeness, with respondents who had completed less than 70% of the survey dropped due to the substantial amount of missing data in their responses. Additionally respondents who reported that they had no experience in child custody evaluation, or who devoted none of their practice to child custody evaluations were also dropped. This resulted in a total of 178 remaining respondents who met selection criteria.

Demographics

The average age of respondents reporting their age (n = 174) was 55.9 years, with minimum and maximum reported ages of respondents of 30 and 77 years, respectively. The standard deviation in age of respondents was 10.15 years. 59% of the sample reported they were female, 41% male. 93.3% identified themselves as Caucasian, 2.8% as Hispanic, 1.7% as

African-American, 1.1% identified as multiracial, and 1.1% of participants did not respond to this question. In regards to their type of location, rural, suburban, or urban, 48.9% of respondents identified their location as suburban, 38.2% as urban, 10.7% as rural, and the remaining 2.2% of respondents did not identify a location type. The vast majority of respondents, 77.5%, reported that they provided their forensic services in private practice, 14.6% reported providing forensic services at a court clinic/agency, 3.4% reported (in a free-form "other" answer option) they provided forensic services in a combination of court clinic/agency and private practice, 2.2% reported providing forensic services at a public mental health clinic, 1.1% at a University Clinic, 0.6% reported other settings, and 0.6% reported no answer to this question.

Degree and Academic Field

44.9% of respondents reported that their highest degree was a Ph.D., 12.4% reported their highest degree was a M.A., 10.7% a Psy.D., 10.1% a MSW, 7.9% a M.S. Each of the following degrees was reported by 3.4% of the respondents: M.D., MSSW, M.Ed., and Other masters level degree. 0.6% of respondents reported their highest degree was an Ed.D. This information is presented below in Table 1.

Table 1 – Highest Degree Reported

Highest Degree	Percent Reporting
Ph.D.	44.9
Psy.D.	10.7
Ed.D	.6
M.D.	3.4
MSW	10.1
MSSW	3.4
M.S.	7.9
M.A.	12.4
M.Ed.	3.4
Other masters level	3.4
Total	100.0

Respondents were asked an open-ended question regarding their academic field of study. Thirteen core academic fields of study were identified, Counseling (including all reported subtypes: Counseling; Counseling and Development; and Marriage and Family Counseling), Social Work (including all reported subtypes: Social Work; Clinical Social Work; and Direct Practice), Psychology (including all reported subtypes: Psychology; Clinical Psychology; Child and Family/General Clinical Psychology; Social-clinical Psychology; Counseling Psychology; Clinical, School, and Community Psychology; School Psychology; School and Community Psychology; Clinical and Counseling Psychology; Counseling Psychology; Child and Clinical Psychology; etc.), Criminal Justice, Human Services, Law, Family Relations/Family Systems, Human Development, Child Development, Sociology, Child and Family, Psychiatry, and Urban and Public Affairs. Percentage distribution is shown in Table 2 below.

Table 2 – Academic Field Reported

Academic Field	Percent Response
Counseling	8.4
Social Work	13.5
Psychology	66.3
Criminal Justice	.6
Human Services	1.1
Law	1.1
Family Relations/Family Systems	2.2
Human Development	.6
Child Development	.6
Sociology	.6
Child and Family	1.1
Psychiatry	2.2
Urban and Public Affairs	.6
No response	1.1

Experience

An interesting issue arose in examining years of experience in mental health, wherein it became clear that some respondents, while actively engaged in child custody evaluation, did not identify as mental health professionals. This included respondents who identified their academic field as Law, Social Work, Human Development, and Urban and Public Affairs, each reporting zero years of experience in mental health. Although their answers were inconsistent with informal expectations that evaluators were mental health professionals, this was not established as a mandatory selection criteria. Rather than repeat intentional exclusion of certain evaluators seen in previous research they have been retained in the dataset.

All 178 respondents reported both their years of experience in mental health and years of experience in child custody evaluation. The mean response regarding years of experience in

mental health was 25.99 years, with a standard deviation of 11.24 years. The mean response regarding years of experience in child custody evaluation as 16.91 years, with a standard deviation of 10.04 years. An aggregate summary of reported training experiences is provided in Table 17 below.

Previous Survey Participation

All 178 respondents reported whether they had ever previously participated in a survey regarding their child custody evaluation practices, with 43.8% (n = 78, 44 masters-level and 34 doctoral-level) reporting that they had never been surveyed in regards to their custody evaluation practices.

Comparison of Groups

One of the initial questions in this research is whether there exist any identifiable differences between masters-level child custody evaluators and their better-studied doctoral-level counterparts. 72 respondents reported masters-level degrees, while the remaining 106 respondents reported doctoral-level degrees.

Demographic Differences

The average reported age of masters-level respondents was 51.58 years, with a standard deviation on 11.45 years. The average reported age of doctoral-level respondents was 58.92 years, with a standard deviation of 7.9 years. An independent-samples t-test was run to determine if there were statistically significant differences in age between the two groups. The assumption

of homogeneity of variances was violated, as assessed by Levene's test for equality of variances (p = 0.000) and thus the Welch t-test was utilized to address this. There was a statistically significant difference in age between masters-level and doctoral-level respondents, with masters-level respondents reporting being younger than doctoral-level respondents by 7.35 years (95% CI, 10.45 to 4.24), $t(114.977) = -4.69$, $p = 0.000$. Further analysis yielded a Cohen's d of -0.747, a medium effect size (between 0.2 and 0.8 in absolute value). Percent of variance explained in respondents' ages by differences in degree was $r^2 = 0.161$, a medium (between 0.09 and 0.25) effect level.

Respondents from both degree types were predominantly Caucasian. A summary of reported ethnicity by degree type is reported in tabular format below.

Table 3 – Respondent race/ethnicity by degree level

	Masters-level	Doctoral-level
African-American	3	0
Caucasian	67	99
Hispanic	1	4
Asian	0	0
Multiracial	1	1
No answer	0	2

One noted difference between the respondents was that there were far fewer male masters-level respondents in the sample, as shown in Table D below. A Chi-square test for association (independence) was conducted between gender and degree level. All expected cell frequencies were greater than five. There was a statistically significant association between

gender and degree level, $\chi^2(1) = 10.686$, $p = 0.001$. The strength of this association was weak, with $\varphi = 0.245$, $p = 0.001$

Table 4 - Respondent gender by degree level

	Masters-level	Doctoral-level
Female	53	52
Male	19	54
No answer	0	0

Regarding location type, 71 masters-level respondents provided information, as did 103 doctoral-level respondents. A Chi-square test for association (independence) was conducted between location and degree level. All expected cell frequencies were greater than five. There was no statistically significant association between location and degree level, $\chi^2(2) = 0.732$, $p = 0.694$.

Table 5 – Location by degree level

	Masters-level	Doctoral-level
Rural	9	10
Suburban	33	54
Urban	29	39
No response	1	3

Finally, in regard to where respondents provided forensic services, 88.7% of doctoral-level respondents and 61.1% of masters-level respondents reported that they provided their forensic services in private practice, 4.7% of doctoral-level respondents and 29.2% of masters-level

respondents reported providing forensic services at a court clinic/agency, 0.9% of doctoral-level respondents and 6.9% of masters-level respondents reported (in a free-form "other" answer option) they provided forensic services in a combination of court clinic/agency and private practice, 1.9% of doctoral-level respondents and 2.8% of masters-level respondents reported providing forensic services at a public mental health clinic, 1.9% of doctoral-level respondents and 0% of masters-level respondents at a University Clinic, 0.9% of doctoral-level respondents and 0% of masters-level respondents reported other settings, and 0.9% of doctoral-level respondents and 0% of masters-level respondents reported no answer to this question. The exact number of respondents in each category is listed in Table F below.

Table 6 – Forensic setting by degree level

	Masters-level	Doctoral-level
Court clinic/agency	21	5
Private practice	44	94
Public mental health clinic	2	2
University clinic	0	2
Combined court and private practice	5	1
Other	0	1
No response	0	1

A Chi-square test for association (independence) was conducted between forensic setting and degree level, however several cells had expected counts less than five. Chi-square analysis should not be used in cases where expected frequency in any cell is less than five (Gravetter & Wallnau, 2004). This breakdown was re-examined using only those respondents who had indicated they provided forensic services solely in a court clinic/agency, or solely in private

practice. The resulting Chi-square test for association (independence) between forensic setting with these limitations and degree level resulted in all expected cell frequencies greater than five. There was a statistically significant association between forensic setting (limited to court clinic/agency and private practice) and degree level, $\chi^2(1) = 21.85$, $p = 0.000$. This was a medium effect size, $\varphi = 0.365$, $p = 0.000$.

Experience Differences

The average reported years of experience of masters-level respondents (n = 72) in child custody evaluations was 11.72 years, with a standard deviation of 8.44 years. The average reported years of experience of doctoral-level respondents (n = 106) in child custody evaluations was 20.43 years, with a standard deviation of 9.52 years. An independent-samples t-test was run to determine if there were statistically significant differences between the two groups in regards to years of experience in child custody evaluations. There was homogeneity of variances as assessed by Levine's test for equality of variances ($p = 0.124$). There was a statistically significant difference in years of experience in child custody evaluation between masters-level and doctoral-level respondents, with masters-level respondents reporting fewer years of experience than doctoral-level respondents by 8.7 years (95% CI, 11.45 to 5.97), $t(176) = -6.269$, $p = 0.000$. An effect size result of $d = 0.965$ was computed for this result, indicating that not only was the difference statistically significant the effect size was large (>0.8 in absolute value). The percent of variance explained by differences in degree level was $r^2 = 0.183$ for years of experience in child custody evaluation, a medium (between 0.09 and 0.25) effect. As can be seen from the means for each group, doctoral level practitioners who responded to this survey have

almost twice the years of experience, on average, as their masters level counterparts, thus these results are not surprising.

Career total child custody evaluations were also reviewed, with masters-level respondents (n = 71) reporting a mean number of 260.17 evaluations, with a standard deviation of 555.59 evaluations. Doctoral-level respondents (n = 105) reported a mean number of 328.86 evaluations, with a standard deviation of 815.96 evaluations. Again an independent-samples t-test was run to determine if there were statistically significant differences between the two groups in regards to total number of child custody evaluations in their careers. There was homogeneity of variances as assessed by Levine's test for equality of variances ($p = 0.641$), however there was no statistically significant difference in total career child custody evaluations between masters-level and doctoral-level respondents $t(174) = -0.619$, $p = 0.537$.

Differences in Percent of Practice Devoted to Evaluation

Masters-level respondents (n = 72) reported on average 49.43% of their practice was devoted to child custody evaluations, with a standard deviation of 36.02%. Doctoral-level respondents (n=106) reported on average 40.69% of their practice devoted to child custody evaluations, with a standard deviation of 29.82%. Using an independent samples t-test to explore whether there were statistically significant differences resulted in a discovery that the assumption of homogeneity of variances was violated, as assessed by Levene's test for equality of variances ($p = 0.015$) and thus the Welch t-test was utilized to address this. There was no statistically significant difference in the percentage of practice devoted to child custody evaluations between masters-level and doctoral-level practitioners, $t(133.006) = -0.619$, $p = 0.091$

Formats for Initiating Involvement and Timeframes

Respondents were asked in what manner they prefer to be brought in to a case as a child custody evaluator. In addition to the historically provided response (retained by one parent/attorney; retained by both parents/attorneys, and appointed by the court or guardian ad litem) two additional "other" responses were repeatedly noted and coded for: "assigned case by agency manager (not individually appointed)" and the combination "either retained by both or appointed by the court." All respondents provided answers to this question, with specific responses shown in the table below.

Table 7 – Format for initiating involvement of the child custody evaluator

	Masters-level	Doctoral-level
Retained by one parent/attorney	0	1
Retained by both parents/attorneys	6	5
Appointed by the court or GAL	58	98
Other	2	0
Assigned case be agency manager	3	0
Either retained by both or appointed by court	3	2
Total	72	106

Most (n = 174) respondents reported information regarding the percentage of child custody evaluations they perform that are court ordered. Masters-level respondents (n = 68) reported on average 90.44% of the child custody evaluations they conduct are court ordered, with a standard deviation of 25.93%. Doctoral-level respondents (n=106) reported on average 93.12% of the child custody evaluations they conduct are court ordered, with a standard deviation of 17.87%. An independent-samples t-test was run to determine if there were statistically significant differences between the two groups in regards to percentage of child custody evaluations that are

court ordered. There was homogeneity of variances as assessed by Levine's test for equality of variances ($p = 0.066$), however there was no statistically significant difference in percentage of child custody evaluations that are court ordered between masters-level and doctoral-level respondents $t(172) = -0.808$, $p = 0.420$.

Finally in regards to formats and time frames, evaluators were asked the time it took from the point of first interview[9] to completion of their report, to complete a typical child custody evaluation. Masters-level respondents (n = 69) reported an average 12.2 weeks to completion, with a standard deviation of 9 weeks. Doctoral-level respondents (n = 101) reported an average 11.7 weeks to completion, with a standard deviation of 5.4 weeks. Using an independent samples t-test to explore whether there were statistically significant differences resulted in a discovery that the assumption of homogeneity of variances was violated, as assessed by Levene's test for equality of variances ($p = 0.013$) and thus the Welch t-test was utilized to address this. There was no statistically significant difference in completion time of custody evaluations between masters-level and doctoral-level practitioners, $t(100.884) = 1.255$, $p = 0.212$.

Differences in Previous Research Participation

61% of masters-level respondents and 32% of doctoral-level respondents reported never having participated in a previous survey regarding their child custody evaluation practices. The exact number of respondents is listed in Table 8 below.

[9] The point of the first interview was selected to eliminate issues such as court procedural time and jurisdictional differences, and better focus on the child custody evaluation process itself.

Table 8 – Previous survey participation by degree level

	Previous survey participation	
	Yes	No
Masters-level	28	44
Doctoral-level	72	34

Hypothesis Findings

The eight major hypotheses initially proposed for this research were evaluated utilizing the data from the respondent sample. Results are presented hypothesis by hypothesis below.

Hypothesis 1

There will be no significant difference between doctoral-level evaluators and masters-level evaluators regarding time spent on interviewing, parent-child observation, and records review.

Hypothesis 1a – Interviewing

The issue of time spent interviewing was examined individually (survey questions 21a, regarding time spent interviewing parents; 21b, time spent interviewing children; 21c time spent interviewing significant others; and 21d other interviews) and as an aggregate (combining these responses). For statistical purposes blank responses to individual items were computed as zero hours spent in that activity. The mean responses and standard deviation are presented in tabular format below.

Table 9 – Mean hours in interviewing

	Masters-level or Doctoral-level	Mean	Std. Deviation
Hours interviewing parents	Masters-level	8.2222	4.71803
	Doctoral-level	9.7547	4.87327
Hours interviewing children	Masters-level	2.9792	2.32021
	Doctoral-level	3.2901	1.79189
Hours interviewing significant others	Masters-level	2.4653	2.78998
	Doctoral-level	2.8608	2.50763
Hours other interviews (outside of collateral contacts)	Masters-level	1.3125	1.63465
	Doctoral-level	1.6132	3.47048
Aggregate hours interviewing	Masters-level	14.9792	8.78637
	Doctoral-level	17.5189	8.47740

For each of these interviewing formats an independent-samples t-test was run to determine if there were statistically significant differences between masters-level and doctoral-level evaluators in regards to time spent interviewing. Levene's test for equality of variances was met for each condition with $p = 0.564$ for time interviewing parents, $p = 0.354$ for time interviewing children, $p = 0.496$ for time interviewing significant others, $p = 0.151$ for time spent in other interviews, and $p = 0.957$ for aggregate time interviewing. The only analysis that showed any statistically significant difference was in regards to time spent interviewing parents, where masters-level respondents reported spending 1.53 (95% CI -2.98 to -0.82) fewer hours in interviewing parents than doctoral-level respondents $t(176) = -2.086$, $p = 0.038$. Further analysis indicated this was a medium effect size, $d = -0.32$; the percent of variance explained by respondents' differences in degree levels was small ($r^2 < 0.09$) , $r^2 = 0.48$.

The non-significant results showed for time interviewing children $t(176) = -1.007$, $p = 0.315$, time interviewing significant others $t(176) = -0.987$, $p = 0.325$, time in other interviews $t(176) = -0.685$, $p = 0.494$, and aggregate time interviewing $t(176) = -1.933$, $p = 0.055$.

Hypothesis 1b – Parent-child observations

The issue of time spent in parent-child observation was examined by asking respondents about the average/typical amount of total time in hours spent in observations in an office/playroom (survey question 21L) and in observations during home visits (survey question 21m). Similar to interviewing time an aggregate time spent in parent-child observations was also calculated by combining these responses. For statistical purposes blank responses to individual items were computed as zero hours spent in that activity. The mean responses and standard deviation are presented in tabular format below.

Table 10 – Mean hours in parent-child observation

	Masters-level or Doctoral-level	Mean	Std. Deviation
Hours of parent-child observation in office/playroom	Masters-level	0.9444	1.28796
	Doctoral-level	2.7799	9.59776
Hours of parent-child observation during home visits	Masters-level	1.9681	1.76992
	Doctoral-level	2.1580	3.27060
Aggregate hours of parent-child observations	Masters-level	2.9125	2.37635
	Doctoral-level	4.9379	12.03540

For each of these parent-child observation formats an independent-samples t-test was run to determine if there were statistically significant differences between masters-level and doctoral-level evaluators in regards to time spent in parent-child observation. Levene's test for

equality of variances was met for parent-child observation in office/playroom, and for the aggregate hours of parent-child observations, with $p = 0.306$ and $p = 0.215$ respectively. The assumption of equality of variances, assessed with Levene's test, was violated in regards to parent-child observation during home visits, with $p = 0.045$ and thus a Welch t-test was utilized rather than a standard t-test with this variable. No statistically significant differences were noted between masters-level and doctoral-level respondents in any observation format, either in office/playroom $t(176) = -1.611$, $p = 0.109$, home visits $t(168.687) = -0.500$, $p = 0.618$, or in the aggregate $t(176) = -1.408$, $p = 0.161$.

Hypothesis 1c – Records review

Time spent in review of records was assessed via survey question 21o, the average/typical amount of total time (in hours) in reviewing materials. The mean response to this question for masters-level evaluators (n = 72) was 6.04 hours, with a standard deviation of 6.31 hours; the mean response for doctoral-level evaluators (n=106) was 7.54 hours with a standard deviation of 7.66 hours. An independent-samples t-test comparing responses revealed the assumption of equality of variances, assessed with Levene's test, was met ($p = 0.257$) but there was no significant difference between the two degree levels in regards to records review, $t(176) = -1.382$, $p = 0.169$.

Hypothesis 2

Doctoral-level evaluators will focus more time on psychometric testing than masters-level evaluators.

Psychometric testing was examined for both adults (survey question 21p) and for children (survey question 21q), as well as in aggregate based on a combination of these responses. For statistical purposes blank responses to individual items were computed as zero hours spent in that activity. The mean responses and standard deviation are presented in tabular format below.

Table 11 – Mean hours of psychometric testing

	Masters-level or Doctoral-level	Mean	Std. Deviation
Hours of psychometric testing of adults	Masters-level	0.5556	1.85297
	Doctoral-level	4.1670	3.03430
Hours of psychometric testing of children	Masters-level	0.2639	1.07459
	Doctoral-level	1.3160	1.99681
Aggregate hours of psychometric testing	Masters-level	0.8194	2.75950
	Doctoral-level	5.4830	4.58134

An independent-samples t-test was run to determine if there were statistically significant differences between masters-level and doctoral-level evaluators in regards to time spent in psychometric testing. The assumption of equality of variances was violated in each condition as assessed by Levene's test for equality of variances ($p = 0.000$ in each condition) and thus again the alternative Welch t-test was utilized. In each condition doctoral-level respondents reported statistically significantly higher amounts of time spent in psychometric testing. Doctoral-level responded reported spending 3.6 (95% CI 4.34 to 2.89) more hours in psychometric testing of

adults, $t(174.263) = -9.846$, $p = 0.000$, and 1.05 (95% CI 1.51 to 0.59) more hours in psychometric testing of children, $t(168.362) = -4.542$, $p = 0.000$, than their masters-level counterparts. The effect sizes of these differences were $d = 1.4$ (large effect), and $d = 0.66$ (medium effect) respectively. Percent of variance accounted for by degree type was $r^2 = 0.36$ for psychometric testing of adults, a large effect, and $r^2 = 0.11$ for psychometric testing of children, a medium effect. Examining aggregate use of psychometric testing, doctoral-level respondents reported spending 4.66 (95% CI 5.75 to 3.58) more hours in psychometric testing in aggregate than their masters-level counterparts, $t(173.796) = -8.462$, $p = 0.000$, $d = 1.2$, $r^2 = 0.29$ (both Cohen's d effect size and percent of variance accounted for indicating large effect sizes).

Hypothesis 3

> Masters-level evaluators will focus more time on naturalistic observation (home visits) than doctoral-level evaluators.

This hypothesis was examined through comparison of respondent responses to questions regarding the average/typical time spend on home visits, other than parent-child observation time, parent-child observations during home visits, and in an aggregate of these two responses. Differences in hours of parent-child observation during home visits were previously reported above in Table 10, and are repeated below for clarity.

Table 12 – Mean hours in home visits

	Masters-level or Doctoral level	Mean	Std. Deviation
Hours spent in home visits (other than parent-child observation)	Masters-level	1.1764	1.89384
	Doctoral-level	0.7028	1.74544
Hours of parent-child observation during home visits	Masters-level	1.9681	1.76992
	Doctoral-level	2.1580	3.27060
Aggregate hours in home visits (both observation and other)	Masters-level	3.1444	3.18393
	Doctoral-level	2.8608	4.34852

For each of these reports of time spent in home visits an independent-samples t-test was run to determine if there were statistically significant differences between masters-level and doctoral-level evaluators in regards to time spent in parent-child observation. Levene's test for equality of variances was met for home visits other than parent-child observations, and for the aggregate hours home visits, with $p = 0.289$ and $p = 0.410$ respectively. The assumption of equality of variances, assessed with Levene's test, was violated in regards to parent-child observation during home visits, with $p = 0.045$ and thus a Welch t-test was utilized rather than a standard t-test with this variable. No statistically significant differences were noted between masters-level and doctoral-level respondents in regards to time spent in home visits, either in time spent in home visits outside of parent-child observation $t(176) = 1.716$, $p = 0.088$, time spent in home visits for parent-child observation $t(168.687) = -0.500$, $p = 0.618$, or in the aggregate $t(176) = 0.474$, $p = 0.636$.

Hypothesis 4

Masters-level evaluators will be more likely to be employed in an agency setting or affiliated directly with a court services office than doctoral-level evaluators.

The demographic breakdown of where respondents provided forensic services is included above in Table F in the demographic differences sub-section, as is a Chi-square test for association (independence) for respondents who had indicated they provide forensic services solely in a court clinic/agency, or solely in private practice. Given that responses to the "forensic setting" are categorical (nominal) data, t-test analysis is not possible as t-tests assume a continuous dependent variable.

Hypothesis 5

Costs for services by masters-level evaluators will be significantly lower than for doctoral-level evaluators.

Costs for services were assessed in multiple ways in the survey in order to examine various factors in how costs were assessed and what those total costs were. In regards to how evaluators charged in a typical case initial survey responses offered were "per hour," "per case," and "other" with an option for extended explanation of what the "other" method of charging. The most common "other" response (n=8) reported that there were no charges to the parties for the evaluator's services; this response was separated out for more detailed assessment. Not all respondents provided information for this question. One responded who reported a "per case"

rate but then reported that there was no charge (their rate = 0) was also recoded to "no charge." A detailed breakdown is outlined in the table below.

Table 13 – Breakdown of how charges are assessed

		Masters-level	Doctoral-level	Total responses
How do you charge in a typical case	Per hour	28	70	98
	Per case	24	24	49
	Other	11	8	19
	No charge	6	2	7
	Total responses	69	104	173

A breakdown of average hourly costs was calculated selecting out only those respondents who noted a "per hour" charge. Two doctoral-level respondents who reported the same "per hour" charge as their total fee for a custody evaluation were also deleted from this analysis as the resulting outliers appeared to be clearly an error (responding to average total fee for an evaluation rather than their hourly rate). This resulted in usable responses from 28 masters-level evaluators and 63 doctoral-level evaluators (not all doctoral-level respondents who reported a per hour charge reported what that charge was).

Table 14 – Mean charges per hour (when charges are per hour)

	Masters-level or Doctoral level	N	Mean	Std. Deviation
Hourly charge in cases where charge is per hour	Masters-level	28	$142.0357	72.05270
	Doctoral-level	63	$239.7619	71.21882

An independent samples t-test was run to determine if there were differences in hourly charges between masters-level and doctoral-level evaluators. There was homogeneity of variances as assessed by Levene's test for equality of variances (p = 0.673). Masters-level evaluators were less expensive than doctoral-level evaluators, a statistically significant difference of $97.73 per hour (95% CI $129.98 to 65.47), $t(89) = -6.020$, p = 000. The difference attributable to degree levels evidenced a large effect size, $d = -1.36$, with the percent of variance explained by respondents' difference in degree levels also indicating a large effect, $r^2 = 0.289$.

A breakdown of average per case costs was calculated selecting out only those respondents who noted a "per case" charge as well. As seen from the table below, not all doctoral-level respondents who noted a per case charge also reported what that fee was.

Table 15 – Mean charges per case (when charges are per case)

	Masters-level or Doctoral-level	N	Mean	Std. Deviation
Per case charges for cases where charge is per case	Masters-level	24	2443.7500	1828.95269
	Doctoral-level	20	5722.5000	3326.38871

An independent samples t-test was run to determine if there were differences in per case charges between masters-level and doctoral-level evaluators. Assumptions regarding homogeneity of variances were violated as assessed by Levene's test for equality of variances (p = 0.017), and thus the Welch t-test was used to address this. Masters-level evaluators were less expensive than doctoral-level evaluators, a statistically significant difference of $3278.75 per case (95% CI 4982.71 to 1574.79), $t(28.296) = -3.940$, p = 000. The difference attributable to

degree levels evidenced a large effect size, d = -1.22, with the percent of variance explained by respondents' difference in degree levels also indicating a large effect, r^2 = 0.35.

Finally in assessing this hypothesis the reported average total fee for a custody evaluation was examined. 162 of the 178 respondents in the data set provided this information, which is broken down in table and chart format below.

Table 16 – Mean total fee for a custody evaluation

	Masters-level or Doctoral level	N	Mean	Std. Deviation
Average total fee for a custody evaluation	Masters-level	65	3217.846	3310.7611
	Doctoral-level	97	8356.186	4997.5372

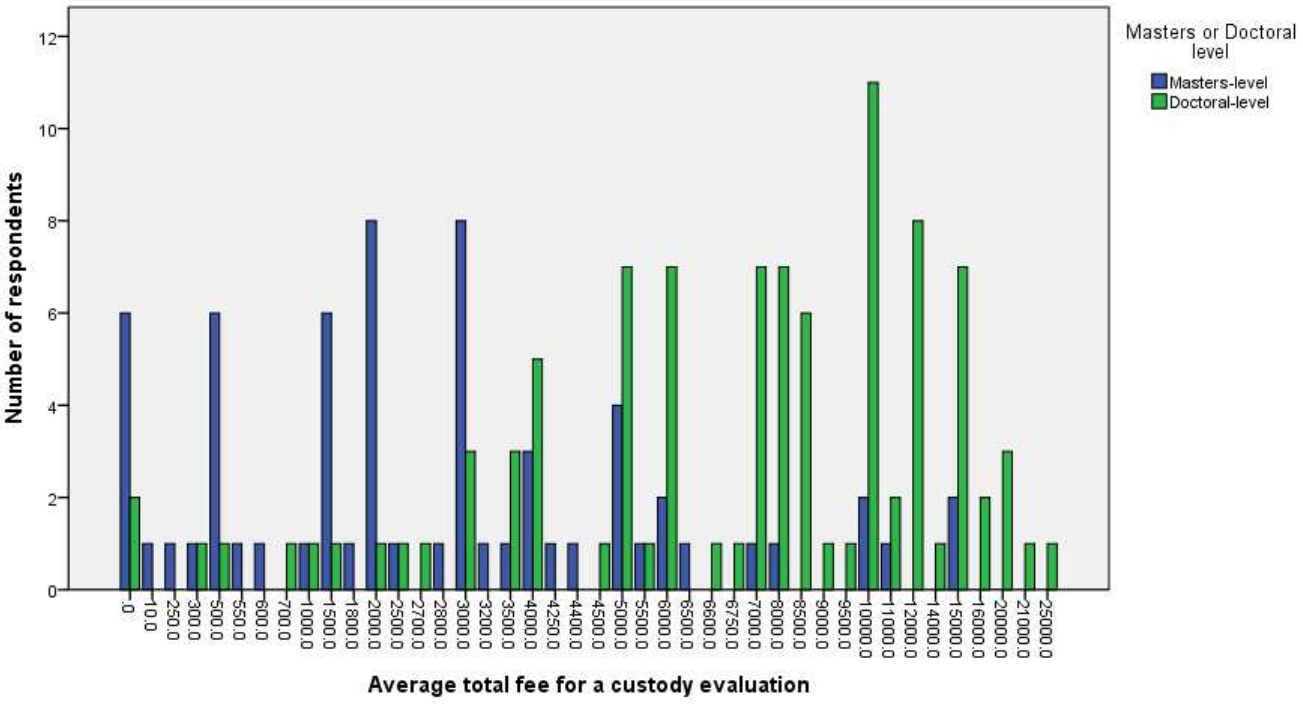

Figure 1 – Bar chart of average total fees

An independent samples t-test was run to determine if there were differences in average total fees for an evaluation between masters-level and doctoral-level respondents. Assumptions regarding homogeneity of variances were violated as assessed by Levene's test for equality of variances (p = 0.001), and thus the Welch t-test was used to address this. Masters-level evaluators were less expensive than doctoral-level evaluators, a statistically significant difference of $5138.34 per case (95% CI 6427.50 to 3849.18), t(159.988) = -7.872, p = 000. The difference attributable to degree levels evidenced a large effect size, d = -1.21, with the percent of variance explained by respondents' difference in degree levels also indicating a large effect, r^2 = 0.279

Hypothesis 6

There will be no significant difference between the specialized training received by either group of evaluators.

Specialized training was assessed by respondent reports as to whether they have received training in child custody evaluations through a graduate course or graduate courses in forensic mental health; a graduate course or graduate courses in assessment of children and families; practicum (in child custody evaluation); internship or post-graduate placement experience, 100% in custody evaluation; internship or post-graduate placement experience, partially in custody evaluation; reading book or journal articles; attending seminars or workshops; post-graduate supervision; research in child custody; or other areas. Common "other" responses included consultation/collaboration with colleagues or peers, on the job training, and alternative dispute resolution training (in mediation and parenting coordination-type services). A tabular

presentation of the absolute number and percentage of respondents (broken down by degree level) reporting of each type of training is included below.

Table 17 – Number of respondents (and percentage) reporting various training

Type of training in child custody evaluation	Masters-level (n = 72)	Doctoral-level (n = 106)	Total (n = 178)
Graduate course(s) in forensic mental health	21 (29%)	31 (29%)	52 (29%)
Graduate course(s) in assessment of children and families	46 (64%)	77 (73%)	123 (69%)
Practicum	12 (17%)	21 (20%)	33 (19%)
Internship/post-graduate placement experience, 100% custody evaluation	3 (4%)	2 (2%)	5 (3%)
Internship/post-graduate placement experience, partial custody evaluation	6 (8%)	23 (22%)	29 (16%)
Books/journal articles	65 (90%)	103 (97%)	168 (94%)
Seminars and workshops	69 (96%)	105 (99%)	174 (98%)
Post-graduate supervision	33 (46%)	62 (58%)	95 (53%)
Research in child custody	29 (40%)	45 (42%)	74 (42%)
Other (all “other” responses)	22 (31%)	17 (16%)	39 (22%)
Other – consultation/collaboration with colleagues/peers	4 (6%)	10 (9%)	14 (8%)
Other – on the job training	9 (13%)	6 (6%)	15 (8%)
Other – alternative dispute resolution training	2 (3%)	1 (1%)	3 (2%)

Chi-square tests for association (independence) were conducted between degree level and the endorsement of various specialty training types, running each as a 2x2 square (degree level and yes/no endorsement of the various trainings). Analysis of "Internship/post-graduate placement experience, 100% custody evaluation," "Seminars and workshops," and "Other – alternative dispute resolution" each resulted in two expected cell frequencies of less than five. Analysis of "Books/journal articles" resulted in one expected cell frequency less than five. Chi-square analysis should not be used in cases where expected frequency in any cell is less than five (Gravetter & Wallnau, 2004). The issues with cell count in the "Seminars and workshops" and "Books/journal articles" training types appear to be due to a high percentage of respondents endorsing having had this type of training, with the issues in the other training types being due to the opposite problem (low endorsement). In such cases the use of Fisher's Exact Test (Fisher, 2000) to determine if there are non-random associations between the variables would be appropriate. Each of these cases resulted in $p > 0.05$, indicating no statistically significant association between the respondent's degree level and having received these types of training. Exact results are reported in the table below.

Table 18 – Fisher's Exact Test results for training types with expected cell frequencies less than five

Type of training in child custody evaluation	Fisher's Exact Significance (2-sided)
Internship/post-graduate placement experience, 100% custody evaluation	$p = 0.395$
Seminars and workshops	$p = 0.305$
Other – alternative dispute resolution	$p = 0.566$
Books/journal articles	$p = 0.092$

In regards to the remaining training types, all expected cell frequencies were greater than five. Chi-square results and significance level are reported for these training types on the table below.

Table 19 – Chi-square results, significance level, and phi coefficient (effect size) for various training received

Type of training in child custody evaluation	χ^2
Graduate course(s) in forensic mental health	$\chi^2(1) = 0.000$ $p = 0.991$
Graduate course(s) in assessment of children and families	$\chi^2(1) = 1.538$ $p = 0.215$
Practicum	$\chi^2(1) = 0.281$ $p = 0.596$
Internship/post-graduate placement experience, partial custody evaluation	$\chi^2(1) = 5.616$ $p = 0.018$
Post-graduate supervision	$\chi^2(1) = 2.760$ $p = 0.097$
Research in child custody	$\chi^2(1) = 0.084$ $p = 0.773$
Other (all “other” responses)	$\chi^2(1) = 5.282$ $p = 0.022$
Other – consultation/collaboration with colleagues/peers	$\chi^2(1) = 0.890$ $p = 0.345$
Other – on the job training	$\chi^2(1) = 2.599$ $p = 0.107$

As the results above indicate, the only significant ($p < 0.05$) associations between degree level and training type were for “Internship/post-graduate placement experience, partial custody evaluation” and the combined “Other” category. In regards to the statistically significant association for “Internship/post-graduate placement experience, partial custody evaluation,” this was a small effect size, $\varphi = 0.178$, $p = 0.018$. In regards to the statistically significant association

for “Other,” this was also a small effect size, $\varphi = -0.172$, $p = 0.022$. Further exploration of internships/post-graduate placement experiences was done by combining responses from both categories of internship/post-graduate placement experience (100% custody evaluation and partial custody evaluation) and again computing a Chi-square test for association (independence) between degree level and participation this type of training when viewed in aggregate. There was no statistically significant association between the combined internship/post-graduate experience variable and degree level, $\chi^2(1) = 3.409$, $p = 0.065$.

Hypothesis 7

Both groups of evaluators will report similar ratings on various decision making variables considered in making recommendations regarding child custody issues.

Respondents were asked to rate the importance of multiple factors in recommending sole or joint legal custody and sole or joint physical custody, using Schepard’s (2004) common definition of those terms. Independent samples t-tests were run to determine if there were differences in ratings of importance of these factors between masters-level and doctoral-level evaluators.

Hypothesis 7a – Legal custody decision making factors

In regards to legal custody the following table lists a breakdown of the number of respondents from each degree level who provided a response regarding each decision making factor as well as statistical information regarding the results obtained:

Table 20 – Legal joint custody factor responses

	Masters-level or Doctoral level	N	Mean	Std. Deviation	T-test results
1. Expressed wishes of the child, age 15	Masters-level	65	6.169	2.1978	$t(157) = -0.613$, $p = 0.541$
	Doctoral-level	94	6.383	2.1357	
2. Expressed wishes of the child, age 10	Masters-level	64	4.594	1.7881	$t(156) = 0.102$, $p = 0.919$
	Doctoral-level	94	4.564	1.8290	
3. Expressed wishes of the child, age 5	Masters-level	65	3.400	1.8097	$t(156) = 2.222$, $p = 0.028$
	Doctoral-level	93	2.785	1.6408	
4. Age of the parents	Masters-level	65	3.123	2.0195	$t(157) = .0630$, $p = 0.529$
	Doctoral-level	94	2.915	2.0669	
5. Marital status of each parent; remarried, single, or cohabiting	Masters-level	65	3.077	2.0639	$t(157) = -0.937$, $p = 0.350$
	Doctoral-level	94	3.394	2.1161	
6. Whether or not one parent is involved in a homosexual relationship	Masters-level	65	2.400	2.0295	$t(157) = 0.619$, $p = 0.537$
	Doctoral-level	94	2.213	1.7590	
7. Number of children in the family	Masters-level	65	3.231	2.1122	$t(157) = -0.352$, $p = 0.725$
	Doctoral-level	94	3.362	2.4312	
8. Age of the children	Masters-level	64	4.969	2.5321	$t(156) = -1.123$, $p = 0.263$
	Doctoral-level	94	5.426	2.4951	
9. The geographic proximity of parental homes	Masters-level	64	5.844	2.4638	$t(156) = -0.781$, $p = 0.436$
	Doctoral-level	94	6.138	2.2317	
10. Economic and physical similarities or differences between parental homes	Masters-level	65	3.969	2.2705	$t(156) = -0.597$, $p = 0.551$
	Doctoral-level	93	4.183	2.1717	
11. Economic stability of the parent	Masters-level	65	5.138	2.4423	$t(156) = -0.827$, $p = 0.410$
	Doctoral-level	93	5.441	2.1288	
12. Whether or not the child is placed in day care while the parent works	Masters-level	65	3.662	2.0485	$t(157) = -0.625$, $p = 0.533$
	Doctoral-level	94	3.872	2.1212	
13. Differences between parental discipline styles	Masters-level	65	5.323	2.4245	$t(156) = -1.165$, $p = 0.245$
	Doctoral-level	93	5.753	2.1652	
14. Each parent's previous involvement in caretaking responsibilities	Masters-level	65	6.400	2.2347	$t(157) = -1.151$, $p = 0.251$
	Doctoral-level	94	6.798	2.0769	
15. The quality of relationship the child has with each parent	Masters-level	65	7.000	2.4367	$t(157) = -1.828$, $p = 0.069$
	Doctoral-level	94	7.638	1.9557	

Table 20 – *Continued*

16. Psychological stability of the parents	Masters-level	65	8.077	1.3841	$t(94.903) = -2.014^{a}$
	Doctoral-level	93	8.462	.8151	$p = 0.047$
17. Ability of the parents to separate their interpersonal difficulties from their parenting decisions	Masters-level	65	8.092	1.4548	$t(96.567) = -1.174^{a}$
	Doctoral-level	94	8.330	.8846	$p = 0.243$
18. The amount of anger and bitterness between the parents	Masters-level	65	7.815	1.5899	$t(157) = -0.429$, $p = 0.669$
	Doctoral-level	94	7.915	1.3252	
19. Whether the child exhibits behavior problems at home or school	Masters-level	64	6.250	2.2537	$t(154) = -0.681$, $p = 0.497$
	Doctoral-level	92	6.478	1.9128	
20. Amount of flexibility in parents' work schedules	Masters-level	65	5.031	2.1137	$t(155) = -1.209$, $p = 0.228$
	Doctoral-level	92	5.446	2.1193	
21. Influences of extended family members (e.g., in-laws and close relatives)	Masters-level	65	5.538	2.0393	$t(157) = -0.044$, $p = 0.965$
	Doctoral-level	94	5.553	2.1130	
22. The parents' willingness to enter joint custody arrangements	Masters-level	65	7.108	2.0926	$t(105.647) = -0.670^{a}$
	Doctoral-level	94	7.309	1.4519	$p = 0.504$
23. Differences between parents' religious beliefs	Masters-level	65	4.123	2.3552	$t(157) = -0.286$, $p = 0.775$
	Doctoral-level	94	4.223	2.0434	
24. Current law (in your state)	Masters-level	64	7.547	1.9915	$t(155) = -0.499$, $p = 0.618$
	Doctoral-level	93	7.710	2.0194	
25. Availability of extended family members	Masters-level	64	4.516	2.1231	$t(155) = -1.879$, $p = 0.062$
	Doctoral-level	93	5.194	2.2854	
26. Problems with the law	Masters-level	65	7.369	1.5569	$t(156) = -0.566$, $p = 0.572$
	Doctoral-level	93	7.516	1.6393	
27. Problems with substance abuse	Masters-level	64	8.094	1.3419	$t(155) = -0.279$, $p = 0.781$
	Doctoral-level	93	8.151	1.1882	
28. Cooperation with previous court orders	Masters-level	65	7.908	1.4761	$t(156) = 0.021$, $p = 0.984$
	Doctoral-level	93	7.903	1.2253	
29. Intelligence of the parents	Masters-level	64	5.000	1.9024	$t(151) = 0.701$, $p = 0.484$
	Doctoral-level	89	4.775	1.9929	
30. Gender of child	Masters-level	63	3.238	2.2122	$t(152) = -0.615$, $p = 0.540$
	Doctoral-level	91	3.462	2.2226	
31. Sexual abuse allegation has been made against one of the parents	Masters-level	65	6.938	2.3041	$t(154) = 0.111$, $p = 0.912$
	Doctoral-level	91	6.901	1.8859	

Table 20 – *Continued*

32. Physical abuse allegation has been made against one of the parents	Masters-level	65	6.923	2.2382	$t(155) = 0.131$, $p = 0.896$
	Doctoral-level	92	6.880	1.8388	
33. One of the parents exhibits better parenting skills than the other	Masters-level	65	6.292	2.3961	$t(114.661) = -2.243$[a]
	Doctoral-level	93	7.086	1.8513	$p = 0.027$
34. The child appears to have a closer emotional bond with one of the parents	Masters-level	65	6.000	2.3651	$t(156) = -2.299$, $p = 0.023$
	Doctoral-level	93	6.806	2.0230	
35. One of the parents is more involved with the child than the other	Masters-level	65	6.185	2.3445	$t(154) = -2.024$, $p = 0.045$
	Doctoral-level	91	6.868	1.8691	

[a] Assumption of equality of variances, assessed with Levene's test, was violated in each of these responses and Welch's t-test was utilized rather than a standard t-test with this variable.

As seen in Table T, there were 5 decision making variables where masters-level and doctoral-level evaluators reported statistically significant differences in their ratings of importance: the expressed wishes of the child, age 5; the psychological stability of the parents; one of the parents exhibits better parenting skills than the other; the child appears to have a closer emotional bond with one of the parents; and one of the parents is more involved with the child than the other. The decision making variable "expressed wishes of the child, age 5" was rated 0.62 (95% CI 0.07 to 1.16) points higher by masters-level evaluators than doctoral-level evaluators. Further analysis indicated this was a medium effect size, $d = 0.36$; the percent of variance explained by respondents' differences in degree levels was small, $r^2 = 0.03$. In regards to the "psychological stability of the parents" variable, this was rated 0.39 (95% CI 0.77 to 0.01) points lower by masters-level evaluators than their doctoral-level counterparts. The effect size of the difference in degree levels was medium, $d = -0.34$; the percent of variance explained by respondents' differences in degree levels was small, $r^2 = 0.04$. For the variable "one of the parents exhibits better parenting skills than the other" masters-level evaluators rated this 0.79

(95% CI 1.49 to 0.09) points lower than doctoral-level evaluators. The effect size of the difference in degree level on this rating was medium, d = -0.37; the percent of variance explained by difference in degree levels was again small, r^2 =0.04. The decision making variable "the child appears to have a closer emotional bond with one of the parents" was rated 0.81 (95% CI 1.46 to 0.09) points lower by masters-level evaluators than doctoral-level evaluators. Further analysis indicated this was a medium effect size, d = -0.37; the percent of variance explained by respondents' differences in degree levels was small, r^2 = 0.03. Finally, in regards to the variable "one of the parents is more involved with the child than the other" masters-level evaluators rated this 0.68 (95% CI 1.35 to 0.02) points lower than doctoral-level evaluators. As with the other decision making factors noted above, the effect size of the difference in degree level on this rating was medium, d = -0.32; the percent of variance explained by difference in degree levels was small, r^2 =0.03.

Hypothesis 7b – Physical custody decision making factors

In regards to physical custody the following table lists a breakdown of the number of respondents from each degree level who provided a response regarding each decision making factor as well as statistical information regarding the results obtained:

Table 21 – Physical joint custody factor responses

	Masters-level or Doctoral level	N	Mean	Std. Deviation	T-test results
1. Expressed wishes of the child, age 15	Masters-level	55	6.727	1.8704	$t(136) = -2.182$, $p = 0.031$
	Doctoral-level	83	7.349	1.4686	
2. Expressed wishes of the child, age 10	Masters-level	55	5.109	1.6741	$t(136) = -0.923$, $p = 0.358$
	Doctoral-level	83	5.361	1.5027	
3. Expressed wishes of the child, age 5	Masters-level	54	3.444	1.8188	$t(136) = 0.496$, $p = 0.621$
	Doctoral-level	84	3.286	1.8470	
4. Age of the parents	Masters-level	54	3.704	1.9965	$t(136) = 1.594$, $p = 0.113$
	Doctoral-level	84	3.131	2.0986	
5. Marital status of each parent; remarried, single, or cohabiting	Masters-level	54	3.852	2.2013	$t(135) = -0.220$, $p = 0.826$
	Doctoral-level	83	3.940	2.3392	
6. Whether or not one parent is involved in a homosexual relationship	Masters-level	53	2.736	2.0583	$t(98.821) = 1.347$ [a]
	Doctoral-level	84	2.274	1.7791	$p = 0.181$
7. Number of children in the family	Masters-level	54	4.074	1.9508	$t(135) = 1.589$, $p = 0.114$
	Doctoral-level	83	3.506	2.1033	
8. Age of the children	Masters-level	53	5.434	1.9563	$t(127.122) = -0.222$ [a]
	Doctoral-level	83	5.518	2.4413	$p = 0.825$
9. The geographic proximity of parental homes	Masters-level	53	6.925	1.9597	$t(94.583) = -0.423$ [a]
	Doctoral-level	83	7.060	1.5953	$p = 0.674$
10. Economic and physical similarities or differences between parental homes	Masters-level	54	4.926	1.9986	$t(135) = -0.342$, $p = 0.733$
	Doctoral-level	83	5.048	2.0713	
11. Economic stability of the parent	Masters-level	54	5.796	1.9268	$t(136) = -1.011$, $p = 0.314$
	Doctoral-level	84	6.107	1.6502	
12. Whether or not the child is placed in day care while the parent works	Masters-level	54	4.833	1.9976	$t(135) = 0.580$, $p = 0.563$
	Doctoral-level	83	4.627	2.0643	
13. Differences between parental discipline styles	Masters-level	54	6.167	1.8708	$t(136) = 0.143$, $p = 0.886$
	Doctoral-level	84	6.119	1.9289	

Table 21 – *Continued*

14. Each parent's previous involvement in caretaking responsibilities	Masters-level	54	6.907	1.3635	$t(136) = -1.300$, $p = 0.196$
	Doctoral-level	84	7.250	1.5972	
15. The quality of relationship the child has with each parent	Masters-level	54	7.463	1.4758	$t(135) = -2.414$, $p = 0.017$
	Doctoral-level	83	8.012	1.1738	
16. Psychological stability of the parents	Masters-level	54	8.278	1.0536	$t(87.455) = -1.345$ [a]
	Doctoral-level	84	8.500	0.7525	$p = 0.182$
17. Ability of the parents to separate their interpersonal difficulties from their parenting decisions	Masters-level	53	8.113	1.1035	$t(135) = -0.763$, $p = 0.447$
	Doctoral-level	84	8.250	0.9679	
18. The amount of anger and bitterness between the parents	Masters-level	54	8.056	1.2196	$t(136) = 0.829$, $p = 0.406$
	Doctoral-level	84	7.869	1.3334	
19. Whether the child exhibits behavior problems at home or school	Masters-level	54	6.500	1.5871	$t(136) = -1.292$, $p = 0.199$
	Doctoral-level	84	6.857	1.5841	
20. Amount of flexibility in parents' work schedules	Masters-level	54	6.019	1.8581	$t(135) = -1.971$, $p = 0.051$
	Doctoral-level	83	6.627	1.7016	
21. Influences of extended family members (e.g., in-laws and close relatives)	Masters-level	54	5.648	1.7500	$t(136) = -0.022$, $p = 0.982$
	Doctoral-level	84	5.655	1.6898	
22. The parents' willingness to enter joint custody arrangements	Masters-level	54	7.370	1.8254	$t(135) = -0.293$, $p = 0.770$
	Doctoral-level	83	7.458	1.6253	
23. Differences between parents' religious beliefs	Masters-level	54	3.981	2.0419	$t(134) = 0.771$, $p = 0.442$
	Doctoral-level	82	3.707	2.0214	
24. Current law (in your state)	Masters-level	53	7.528	2.0995	$t(134) = -0.361$, $p = 0.719$
	Doctoral-level	83	7.663	2.1258	
25. Availability of extended family members	Masters-level	54	5.222	2.0619	$t(134) = -1.543$, $p = 0.125$
	Doctoral-level	82	5.756	1.9152	
26. Problems with the law	Masters-level	54	7.741	1.4033	$t(136) = -0.569$, $p = 0.570$
	Doctoral-level	84	7.881	1.4177	
27. Problems with substance abuse	Masters-level	54	8.352	0.9144	$t(136) = 0.658$, $p = 0.512$
	Doctoral-level	84	8.238	1.0369	
28. Cooperation with previous court orders	Masters-level	52	8.019	1.1113	$t(134) = -0.021$, $p = 0.983$
	Doctoral-level	84	8.024	1.3078	

Table 21 – *Continued*

29. Intelligence of the parents	Masters-level	53	4.925	1.8171	$t(117.547) = 0.628$[a]
	Doctoral-level	81	4.716	1.9700	$p = 0.531$
30. Gender of child	Masters-level	53	3.377	2.3715	$t(132) = -0.614$, $p = 0.540$
	Doctoral-level	81	3.630	2.2937	
31. Sexual abuse allegation has been made against one of the parents	Masters-level	54	7.185	2.2240	$t(134) = 0.503$, $p = 0.616$
	Doctoral-level	82	6.988	2.2470	
32. Physical abuse allegation has been made against one of the parents	Masters-level	54	7.185	2.2325	$t(133) = 0.411$, $p = 0.682$
	Doctoral-level	81	7.025	2.2191	
33. One of the parents exhibits better parenting skills than the other	Masters-level	54	6.833	1.7883	$t(134) = -2.541$, $p = 0.012$
	Doctoral-level	82	7.561	1.5243	
34. The child appears to have a closer emotional bond with one of the parents	Masters-level	54	6.815	1.6943	$t(136) = -3.386$, $p = 0.001$
	Doctoral-level	84	7.655	1.2173	
35. One of the parents is more involved with the child than the other	Masters-level	53	6.774	1.6248	$t(135) = -3.266$, $p = 0.001$
	Doctoral-level	84	7.560	1.1858	

[a] Assumption of equality of variances, assessed with Levene's test, was violated in each of these responses and Welch's t-test was utilized rather than a standard t-test with this variable.

As seen in Table 21, there were again 5 decision making variables where masters-level and doctoral-level evaluators reported statistically significant differences in their ratings of importance: the expressed wishes of the child, age 15, the quality of relationship the child has with each parent, one of the parents exhibits better parenting skills than the other, the child appears to have a closer emotional bond with one of the parents, and one of the parents is more involved with the child than the other. The decision making variable "expressed wishes of the

child, age 15" was rated 0.62 (95% CI 1.19 to 0.06) points lower by masters-level evaluators than doctoral-level evaluators. Further analysis indicated this was a medium effect size, d = -0.37; the percent of variance explained by respondents' differences in degree levels was small, r^2 = 0.03. In regards to the "the quality of relationship the child has with each parent" variable, this was rated 0.55 (95% CI 1.0 to 0.10) points lower by masters-level evaluators than their doctoral-level counterparts. The effect size of the difference in degree levels was medium, d = -0.41; the percent of variance explained by respondents' differences in degree levels was small, r^2 = 0.04. For the variable "one of the parents exhibits better parenting skills than the other" masters-level evaluators rated this 0.73 (95% CI 1.29 to 0.16) points lower than doctoral-level evaluators. The effect size of the difference in degree level on this rating was medium, d = -0.44; the percent of variance explained by difference in degree levels was again small, r^2 = 0.05. The decision making variable "the child appears to have a closer emotional bond with one of the parents" was rated 0.84 (95% CI 1.33 to 0.35) points lower by masters-level evaluators than doctoral-level evaluators. Further analysis indicated this was a medium effect size, d = -0.57; the percent of variance explained by respondents' differences in degree levels was small, r^2 = 0.08. Finally, in regards to the variable "one of the parents is more involved with the child than the other" masters-level evaluators rated this 0.79 (95% CI 1.26 to 0.31) points lower than doctoral-level evaluators. As with the other decision making factors noted above, the effect size of the difference in degree level on this rating was medium, d = -0.55; the percent of variance explained by difference in degree levels was small, r^2 =0.07.

Hypothesis 8

More experienced evaluators will spend more time in assessment than less experienced evaluators, regardless of degree level.

In order to examine different levels of experience the pool of respondents were first broken up into three approximately equal groups based percentile breakdown of their years of experience in child custody evaluations: less than 12 years experience (n = 58), 12 to 20 years experience (n = 59), and 21 years experience or more (n = 61). A breakdown of experience levels in child custody evaluation by degree level is listed below.

Table 22 – Child custody evaluation experience separated by degree level

		Masters-level	Doctoral-level
Levels of CCE experience	Low	37 (51%)	21 (20%)
	Medium	25 (35%)	34 (32%)
	High	10 (14%)	51 (48%)
	Total	72	106

While this breakdown allows examination of the effects of general experience it results in an unbalanced design for a one-way ANOVA when further subdivided by degree level, the disparity in ages between masters-level and doctoral-level evaluators rendered attempts to better balance the groups unfeasible. As noted previously, there is a statistically significant difference in age between masters-level and doctoral-level respondents, with masters-level respondents reporting being younger than doctoral-level respondents by 7.35 years (95% CI, 10.45 to 4.24),

$t(114.977) = -4.69$, $p = 0.000$. Time in assessment was calculated by adding the total time in various areas encompassed by survey question 21, with the exception of testifying in court.

When these experience levels (regardless of degree level) were analyzed statistically the assumption of homogeneity of variances was violated, as assessed by Levene's Test of Homogeneity of Variance ($p = 0.009$). As a result a Welch ANOVA was utilized instead of a standard one-way ANOVA. The number of hours spent in assessment was statistically significantly different between the three experience levels, Welch's $F(2, 104.378) = 15.295$, $p = 0.000$. The number of hours spent in assessment increased from the low experience (M = 42.28, SD = 16.8), to moderate experience (M = 59.91, SD = 31.73), to high experience groups (M = 66.37, SD = 35.26), in that order. Games-Howell post-hoc analysis revealed that the mean increase from low experience to moderate experience (17.6, 95% CI 6.5 to 28.8) was statistically significant ($p = 0.001$), however the increase from moderate experience to high experience was not ($p = 0.546$).

When these experience levels were analyzed only for masters-level respondents the assumption of homogeneity of variances was again violated, as assessed by Levene's Test of Homogeneity of Variance ($p = 0.000$). As a result a Welch ANOVA was utilized instead of a standard one-way ANOVA. The number of hours spent in assessment was statistically significantly different between the three experience levels, Welch's $F(2, 17.44) = 5.347$, $p = 0.015$. The number of hours spent in assessment increased from the masters-level evaluators with low experience (M = 38.05, SD = 13.76), to masters-level evaluators with moderate experience (M = 56.52, SD = 27.66), to masters-level evaluators with high experience groups (M = 62.56, SD = 52.25), in that order. Games-Howell post-hoc analysis revealed that the mean increase

from low experience to moderate experience (18.47, 95% CI 3.8 to 33.2) was statistically significant ($p = 0.011$), however the increase from masters-level evaluators with moderate experience to masters-level evaluators with high experience was not ($p = 0.942$).

When these experience levels were analyzed only for doctoral-level respondents there was homogeneity of variances as assessed by Levene's Test of Homogeneity of Variance ($p = 0.458$). The number of hours spent in assessment was not statistically significantly different between the three experience levels, $F(2, 103) = 2.338$, $p = 0.102$.

Finally for the one-way ANOVA testing, all six experience levels (masters-level low, medium, and high; doctoral-level low, medium, and high) were run. The assumption of homogeneity of variances was again violated, as assessed by Levene's Test of Homogeneity of Variance ($p = 0.010$). As a result a Welch ANOVA was utilized instead of a standard one-way ANOVA. The number of hours spent in assessment was statistically significantly different between the six groups, Welch's $F(5, 48.644) = 8.796$, $p = 0.000$. Games-Howell post-hoc analysis revealed that the only statistically significant differences were with low experience masters-level respondents who had lower total assessment hours than the following groups: medium-experience masters-level respondents (18.47, 95% CI 0.4 to 36.6; $p = 0.043$), medium-experience doctoral-level respondents (24.3, 95% CI 5.4 to 43.3; $p = 0.005$), and high-experience doctoral-level respondents (28.99, 95% CI 14.3 to 43.7; $p = 0.000$).

Additional Analysis

Given the unexpected outcomes on hypothesis 1 and 3 regarding interviewing and home visits, combined with the expected outcome of hypothesis 2 regarding psychometric testing it

became clear that an exploration of aggregate hours spent in the various assessment areas (interviews, collateral contacts, observation, home visits, review of materials, psychometric testing, report writing, and meetings/consultation) should be undertaken to see if there were any significant differences between the degree levels. Additionally the unexamined sub areas (collateral contacts, report writing, and meetings/consultation) should also be examined. Those results are presented in tabular format below.

Table 23 -Total hours in assessment, collateral contacts, and meetings to review/attorney consultation

	Masters-level or Doctoral level	Mean	Std. Deviation
Aggregate hours in all assessment areas	Masters-level	47.0007	28.24404
	Doctoral-level	62.1228	31.29120
Total hours collateral contacts	Masters-level	8.2903	7.18050
	Doctoral-level	7.4896	5.39452
Total hours of meetings to review/attorney consultation	Masters-level	1.8993	2.19821
	Doctoral-level	2.6604	3.66518

Independent samples t-tests were run to determine statistically significant differences between the degree-level groups based on the above listed areas. There was homogeneity of variances as assessed by Levine's test for equality of variances in each case ($p = 0.527$, $p = 0.067$, and $p = 0.441$ respectively). The only statistically significant difference was in the aggregate hours in all assessment areas, where masters-level evaluators reported spending less total time by 15.1 hours (95% CI 6.1 to 24.2), $t(176) = -3.290$, $p = 0.001$. An effect size result of $d = -0.51$ was computed for this result, indicating a medium effect size. The percent of variance

explained by differences in degree level was $r^2 = 0.058$ for aggregate hours spent on the evaluation, a small effect.

Results

Overall results regarding hypothesis testing presented a complex and interesting picture of the data.

In regards to hypothesis 1, regarding interview time, parent-child observation, and records review, a statistically significant but small difference in time spent in interviews of parents was observed, with doctoral-level evaluators spending more time interviewing parents than masters-level evaluators. Other than this difference there were no other significant differences noted. Hypothesis 1 is rejected due to the single statistical exception noted.

Analysis of hypothesis 2, regarding use of psychometric testing, as expected yielded clear data that doctoral-level respondents reported significantly higher use of psychometric testing with adults, children, and in the aggregate than their masters-level counterparts. We accept hypothesis 2.

In hypothesis 3 it was expected that masters-level evaluators would focus more on home visits, however this was not supported in the data. No statistically significant differences were found between the degree levels and thus we reject hypothesis 3.

With hypothesis 4 it was believed that employment in an agency or court-affiliated setting would be more likely for masters-level evaluators. There was a statistically significant association between forensic setting (court clinic/agency and private practice) and degree level.

A far higher percentage of masters-level evaluators were found in the court clinic/agency than doctoral-level evaluators. Hypothesis 4 is accepted.

Hypothesis 5 posited that costs for services by masters-level evaluators would be significantly lower than for doctoral-level counterparts. In examining evaluators who charged a per hour fee masters-level evaluators were statistically significantly less expensive; as was also the case for evaluators who charged a per case fee. In examining reports of total fees masters-level evaluators were also statistically significantly less expensive. In all of these situations computed effect sizes were large. Hypothesis 5 is accepted.

In regards to hypothesis 6, no differences between specialized training received were expected. One defined training type (Internship/post-graduate placement experience, partial custody evaluation) demonstrated a statistically significant association with degree level, indicating that there was such a difference. Additionally the "other" category also showed statistically significant association with degree level, again indicating a difference. We reject hypothesis 6 based on these results.

In hypothesis 7 it was expected there would be no reported differences on decision making variables considered in making recommendations on child custody. This was examined in terms of both legal custody decision making factors and physical custody decision making factors, each of which indicated there were five decision making variables where masters-level and doctoral-level evaluators reported statistically significant differences in their ratings of importance (although these areas were different for legal and physical custody). Thus hypothesis 7 is also rejected.

Finally, in hypothesis 8 it was postulated more experienced evaluators would spend more time in assessment than less experienced evaluators, regardless of degree level. While it initially appeared that this was true when the respondents were examined in aggregate (that is based solely on experience level and time in assessment), when doctoral-level evaluators were examined apart from masers-level evaluators no differences statistically significant differences in assessment time were noted in the experience levels coded for. Hypothesis 8 is also rejected.

Table 24 – Comparison of custody evaluator procedures

Procedure	Mean hours spent in procedure (% of respondents using this procedure, where reported)						
	Keilin & Bloom (1986)[10]	Ackerman & Ackerman (1997)	LaFortune & Carpenter (1998)	Gourley & Stolberg (2000)[11]	Bow & Quinnell (2001)	Ackerman & Brey Pritzl (2011)	Robb (2013)
Interviewing parents	4.07 (100%)	4.7	4.85	5.3	7.0+[12]	7.1	9.14
Interviewing children	1.55 (98.8)	2.7	1.44-1.89[13]	3.5	1.75	3.6	3.2
Psychometric testing of adults	2.66 (75.6%)	--	--	--	3.03 (91%)	--	2.92
Psychometric testing of children	2.50 (74.4%)	--	--	--	1.97 (61%)	--	1.02
--Combined adult and child testing	--	5.2	--[14]	4.0	--	6.1	3.94
Parent-child observation	1.89 (68.8%)	2.6	2.49	3.0	1.59 (92%)	3.7	4.1
Interviewing significant others[15]	1.32 (48.8%)	1.6	1.37	2.1	--	2.3	2.73

[10] Respondents who reported not using a particular procedure were not factored in to this calculation.
[11] Data on procedures was only reported from the sample of evaluators rated as "credible" rather than a general sample of evaluators
[12] Like Keilin & Bloom this study examined the percentage of respondents using particular procedures, however it is unclear how they aggregated this information to get the "+" (apparently meaning 7 hours and more) in this figure.
[13] Dependent on age
[14] LaFortune & Carpenter present data on testing on a test by test basis, but it is unclear from their figures the mean hours for all testing procedures

Table 24 – *Continued*

Collateral contacts[16]	--	1.6	1.16	2.2	--[17]	3.2	7.8
Home visits	1.42 (30.0%)	--	--	--	2.14 (33%)	--	1.1[18]
Reviewing materials	--	2.6	1.11[19]	2.4	2.97 (98%)	5.6	6.9
Report writing	2.84 (93.8%)	5.3	--	4.3	7.3	10.6	13.8
Consulting with attorneys	1.36 (91.4%)	1.2	--	1.9	1.31 (28%)	1.3	0.86
Court testimony	2.29 (79.0%)	2.2	--	2.5	--	2.6	2.7

[15] "Significant others" is not adequately operationalized in several of the studies – in Keilin & Bloom (1986), for instance, it is parenthetically noted that this means "friends and relatives" (p.340) whereas no such clarification is provided in later studies. It is unclear how this differs in some studies from "collateral contacts" although it appears to mean spouses and paramours in later studies.

[16] This term is not consistently defined, although various studies focused on interviews and/or questionnaires used with third parties who may have relevant information. It is generally used to mean friends, non-nuclear family members, or professional contacts in line with prevailing nomenclature. (Austin, 2002).

[17] Bow & Quinnell break down collateral contact by several sub-groups: therapists, teachers, physicians, neighbors and friends, and relatives but do not provide a clear cumulative figure.

[18] Excluding parent-child observation during home visits

[19] Limited to school records

Table 25 Comparison of evaluator demographics and experience

Respondent variable	Keilin & Bloom (1986)[20]	Ackerman & Ackerman (1997)[21]	LaFortune & Carpenter (1998)	Gourley & Stolberg (2000)[22]	Gourley & Stolberg (2000)[23]	Bow & Quinnell (2001)[24]	Ackerman & Brey Pritzl (2011)[25]	Robb (2013)
Number of respondents	82	201	165	65	21	198	213[26]	178
Average age (years)	47.7	49.1	--	--	--	51 years	56.2	55.9
Gender: Male / Female	64 / 18	139 / 62	97 / 68	--	17 / 4	103 / 95	130/83	73/105
Average lifetime # of evaluations completed	156.5	214.9	98.2	--	--	245	269	301
Average number of evaluations in last year	16.9	--	8.1	2.47	--	--	--	11.73
Education:								
Doctoral level	96.4%	100%	92.8%	--	--	96%	100%	59.6%
Masters level	3.6%	0%	7.2%	--	--	4%	0%	40.4%
Profession/Academic Field:								
Psychology	78.1%	100%	89%	100%	100%	100%	100%	66.3%
Social work	1.2%	0%	3%	0%	0%	0%	0%	13.5%
Counseling	2.4%[27]	0%	6%	0%	0%	0%	0%	8.4%[28]
Marriage & family therapy	--	0%	1%	0%	0%	0%	0%	--
Psychiatry	18.3%	0%	0%	0%	0%	0%	0%	2.2%

Table 25 – *Continued*

[20] Inclusion criteria required involvement in at least 4 custody evaluations
[21] Inclusion criteria required involvement in at least 10 custody evaluations
[22] Using a general sample
[23] Using a sample of evaluators rated as "credible" rather than a general sample of evaluators
[24] This sample is identical to the companion paper by Quinnell & Bow (2001).
[25] Minimum number of evaluations completed to be included was 10 custody evaluations
[26] Although the published article notes these were "filled out completely" this conflicts with the underlying dissertation which the article is based on, which notes 213 responses used, although not all 213 were fully completed.
[27] Listed as "masters level practitioners" in the study
[28] Combines "Marriage and Family Counseling," "Counseling and Development," and "Counseling" responses

Respondent variable	Keilin & Bloom (1986)	Ackerman & Ackerman (1997)	LaFortune & Carpenter (1998)	Gourley & Stolberg (2000)[29]	Gourley & Stolberg (2000)	Bow & Quinnell (2001)	Ackerman & Pritzl (2011)	Robb (2013)
Experience:								
Average years in practice	16.1	19.0	17.9	--	--	22.66	24.8	25.99[30]
Years in general forensics	--	--	--	--	--	15.62	--	--
Years in child custody field	10.6	--	9.6	9.8	--	13.57	--	16.91
Primary training with both adults and children	--	--	87.3%	--	--	76%	--	--
Training in child custody derived from:								
--Seminars	--	--	--	43.0%	57.1%	86%	95%	98%
--Supervision	--	--	72.1%	36.9%	28.6%	44%	--	53%
--Internship or postdoc	--	--	30.3%	12.3%	14.3%	39%	13.3%	19%[31]
--Graduate forensic course	--	--	21.2%	13.8%	33%	18%	12.8%	29.2%
--Reading books/journals on custody evaluation	--	--	--	78.5%	76.2%	--	--	94%
--Research in the field	--	--	--	[32]	23.8%	--	--	42%
--Other training	--	--	--	21.5%	--	16%	4.3%[33]	22%

[29] Again, this is using a general sample; the other column is a sample of evaluators rated as "credible" rather than a general sample of evaluators
[30] Sample including evaluators who reported no years of mental health practice
[31] Either 100% child custody evaluation or partial child custody evaluation focused.
[32] The study combined this data into the "other" field along with "clinical" experience and other undisclosed factors. It is unclear why research was separated out in the "credible" sample but no comparison with the general sample was given.
[33] Over 300 practicum hours

Table 26 - Breakdown of historical comparison areas from this study

	Masters-level	Doctoral-level	Total
Number of respondents	72	106	178
Average age (years)	51.58	58.92	55.9
Gender: Male / Female	19 / 53	54 / 52	73/105
Average lifetime # of evaluations completed	260	329	301
Average number of evaluations in last year	15.84	8.89	11.73
Profession/Academic Field:			
Psychology	37.5%	85.8%	66.3%
Social work	33.3%	0%	13.5%
Counseling	16.6%	2.8%	8.4%[34]
Psychiatry	0%	3.8%	2.2%
Experience:			
Average years in practice (mental health)	19.9	30.1	25.99[35]
Years in child custody field	11.72	20.43	16.91
Training in child custody derived from:			
--Seminars	96%	99%	98%
--Supervision	46%	58%	53%
--Internship or postdoc	12%	24%	19%[36]
--Graduate forensic course	29%	31%	29.2%
--Reading books/journals on custody evaluation	90%	97%	94%
--Research in the field	40%	42%	42%
--Other training	31%	16%	22%

[continued next page]

[34] Combines "Marriage and Family Counseling," "Counseling and Development," and "Counseling" responses

[35] Sample including evaluators who reported no years of mental health practice

[36] Either 100% child custody evaluation or partial child custody evaluation focused.

Table 26 – *Continued*

Procedure (in hours):	Masters-level	Doctoral-level	Total
Interviewing parents	8.22	9.75	9.14
Interviewing children	2.98	3.29	3.2
Psychometric testing of adults	0.56	4.17	2.92
Psychometric testing of children	0.26	1.32	1.02
--Combined adult and child testing	0.82	5.48[37]	3.94
Parent-child observation	2.91	4.94	4.1
Interviewing significant others	2.47	2.86	2.73
Collateral contacts	8.29	7.49	7.8
Home visits (outside of parent-child observation)	1.18	0.7	1.1[38]
Reviewing materials	6.4	7.54	6.9
Report writing	10.89	15.79	13.8
Consulting with attorneys	0.58	0.85	0.86
Court testimony	1.9	3.3	2.7

[37] Exact addition from preceding columns is off due to rounding
[38] Excluding parent-child observation during home visits

Chapter 6:

Conclusions and Discussion

This study set out to answer some of the pervasive questions regarding the training and preparation of child custody evaluators, particularly how masters-level child custody evaluators practice and if that practice differs from either current practice of doctoral-level evaluators or the previously studied populations of child custody evaluators where investigative samples have primarily been drawn from doctoral-level practitioners. A non-experimental survey approach was utilized to obtain exploratory and descriptive data. In order to obtain a large number of cases from a group with no official listing or compilation of membership a non-probability sample was necessary, and snowball sampling was utilized in order to reach the broadest section of the population of child custody evaluators possible. Targets of 64 members in each group (masters-level and doctoral-level) were set based on projected data analysis. These targets were exceeded in the final sample. It appears clear from the analysis conducted that there are multiple differences between masters-level child custody evaluators and their doctoral-level counterparts. Some of these differences are obvious and expected, given the historical differences in the focus of training, while others are surprising and not altogether clear in what the full interpretation of their real-world significance might be.

Interpretation of Findings

The basic practice information obtained in this study provides an insightful look at the broader population of child custody evaluators in comparison to previous studies. As seen in

Table W, evaluators in this sample reported more hours of interviewing of parents, on average, than evaluators in any previous study, and roughly similar amounts of time interviewing children, however evaluators in the current sample reported less aggregate time in psychometric testing and more than twice the aggregate number of hours in collateral contacts than any other previous survey. Respondents were somewhat older, on average, than respondents in most other surveys (save for in Ackerman and Brey Pritzl, 2011), but had more experience in completing child custody evaluations than previously studied groups as well. Ultimately though the greatest demographic result obtained was in clearly demonstrating that when the sampling frame involves all child custody evaluators over 40% of respondents were masters-level practitioners. This supports previous indirect data, such as reports from family law judges in Bow & Quinnell (2004), on the percentage of child custody evaluations being carried out by masers-level evaluators.

An interesting early finding was in the small number of evaluators who did not self-identify as mental health professionals (reporting no experience in mental health), despite experience in conducting child custody evaluation. It is unclear if this is due to idiosyncratic jurisdictional issues or a poor operationalization of “mental health services” as a concept, however it is suspected, based on the academic field reports, that this may be a reflection of a bifurcation in clinical treatment focus counseling/therapy provision and assessment of family functioning for forensic purposes. Overall implications are unclear as to whether this is merely an artifact of attempting to gather data from non-doctoral-level evaluators or a more subtle differentiation in types of services (or even the conceptualization of the nature of child custody evaluation) across the various jurisdictions where respondents practice.

The marked differences in gender parity in the different degree levels appears to be reflective of a broader societal issue where large percentages of social workers (81.6%) and counselors (69.9%) are female, although this reflection of population percentages does not hold true for doctoral-level evaluators where, for instance, most psychologists (72.7%) are also female (Bureau of Labor Statistics, 2012). Also of interest was the disparity in years of experience between the two degree level groups. Doctoral-level respondents were both significantly older than their masters-level counterparts, and significantly more experienced. While some difference in age might be expected due to differences in length of academic programs this may be an indication of a shift in the population of custody evaluators. Psychologists and psychiatrists were, historically, the first mental health professionals the legal system turned to for aid; however like many mental health services masters-level practitioners may now be providing many of the same services that were traditionally dominated by doctoral-level practitioners. Further study of this issue is warranted.

In regards to issues of objectivity in assessment one of the first ways an evaluator can demonstrate neutrality and a lack of bias to those who are being evaluated is in serving only when they are appointed by a third party or jointly agreed upon by the litigants themselves. This is in contrast to the historically criticized “hired gun” mentality where the provider would be seen, often rightly as beholden to the side that retained them (Hagan, 1997; Murrie, Boccaccinin, Guarenra & Rufino, 2013). It seems clear that evaluators of both degree levels have adopted a neutral starting approach in how they are brought in to cases.

The results of the various hypotheses that were tested indicate clear, meaningful differences in many ways between the two studied degree levels of evaluators. While it is

unsurprising that doctoral-level evaluators focus more time in psychometric testing, given the preponderance of psychologists in that sample, doctoral-level evaluators also spent more time interviewing parents than masters-level evaluators, although the effect size of this difference was small. Also, there was no significant difference between the degree levels regarding naturalistic observation. Additional analysis of hours spent in various assessment areas revealed that when taken in aggregate doctoral-level evaluators spent more time on evaluations than their masters-level counterparts. This difference in time spent in evaluations may explain in part the lower cost for masters-level evaluators, although it fails to shed light on why even when rates are assessed per hour masters-level evaluators are less expensive, per hour, than their doctoral level counterparts. Clearly in regards to prices, doctoral-level evaluators are able to place a premium on their services due to their higher level of educational attainment, however it is unclear, given the minimal statistical differences in regards to specialized training, whether this premium is warranted or not. It seems likely an examination and study of actual work product differences would be necessary to address that question.

There were also surprising differences in regards to the importance of decision making variables. As both groups rated the expressed wishes of a five-year-old in the unimportant to very unimportant range of the Likert-like scale (logical, as children of that age lack any substantive ability to understand complex legal issues) the first factor may be an area where there is a distinction without a practical difference, despite the statistically significant variation between the groups. Likewise both groups rated psychological stability in the very important range of the Likert-like scale, issues of parental involvement, and the child's emotional bond in the somewhat important range of the Likert-like scale, and parental skills in the somewhat

important to important range. Caution is urged in not over-interpreting these differences, especially when there may be other issues, just as jurisdictional differences in legal joint custody considerations in play. In regards to physical joint custody, we see a repeat of three factors from legal joint custody where there are statistically significant differences (parental skills, parental involvement, and child's emotional bond) and again a clustering of the means on each between somewhat important and important for masters-level evaluators and important and very important for doctoral-level evaluators. The two different divergent factors between physical and legal custody responses are expressed wishes of a fifteen-year-old and quality of the relationship the child has with each parent. The first was in the important to somewhat important range, and the second in the important to very important range. Again, caution is urged in interpretation of these issues given how similar the general preferences are. At best it appears that masters-level and doctoral-level respondents differ in minor degrees, rather than in a more glaring manner.

In regards to the final hypothesis evaluated in this study, it appears that masters-level evaluators in the lowest experience group account for differences seen between the various experience levels and the amount of time spent in assessment. It was intriguing to see that when doctoral-level evaluators were separated out there were no effects of experience on the amount of time spent in evaluation. It is entirely possible that this reflects a greater level of initial training in this area as part of their course of academic study, and/or that the effects of increased skills gained in actual practice manifest themselves as more efficient approaches to interviewing rather than a focus on more detailed/lengthy interviewing. It seems likely that whatever benefits experience has brought to evaluator skill level they are not well assessed in the current survey instrument.

Limitations

There are limitations to generalizing the findings of this study. The participants were gathered using snowball sampling. In combination with a lack of ability to establish population parameters for child custody evaluators there is no way to establish the representativeness of the sample obtained. Data collection occurred during a limited time frame and it may be possible that a longer data collection window would produce different results. The use of self reported data is a limitation as there are respondent biases toward underreporting sensitive information as well as limitations on recall. No verification of information was obtained. Data were gathered utilizing a commercially available internet based survey platform, which may have precluded some respondents from participating or led to under participation by respondents uncomfortable with that format. It also appears that many of the respondents to the survey are involved in research in child custody evaluation, which may be an artifact of the snowball sampling which included many of this researcher's professional contacts in the field. A better funded recruitment process and more widely sourced respondents might provide additional depth of information. More studies, both qualitative and quantitative, are necessary to better understand the issues of evaluator practices and preparation.

Implications

There appear to be several potential implications for practice and policy in regards to child custody evaluators based on the data presented. There are clearly multiple implications that should be considered for future research in regards to child custody evaluators.

Practice

One obvious take-away for child custody evaluators, at either the masters-level or doctoral-level, from this research is that if one has not participated in seminars as well as read books or journal articles on child custody evaluation one is well out of step with current training practices. This may seem straightforward to the point of being simplistic, but while a decade ago it might have been argued that books and journals were the only clear consensus method of specialized training this is now the second study to show over 95% of respondents having participated in such seminars. It also appears that there are clear implications in regards to the amount of collateral information that evaluators are obtaining. The number of hours of collateral contacts reported in this study is more than double the number of hours reported in any previous study on evaluators, and far outstrips the amount of time spent in psychometric testing. While this study will certainly add to the debate on the utility of psychometric testing, it seems that calls for obtaining more "real-world" behavioral documentation (e.g. Austin, 2002) are seeing fruition in practice. It also appears that this study further highlights the possible trend toward more time being invested in parent-child observation as well. Evaluators are well served in being aware of these issues, and in understanding how their practices are similar or different from these issues, so that when challenged in court they will be able to speak to where their methods falls on the broad spectrum of approaches to child custody evaluation.

Policy

It also appears clear that while professionals with an academic background in psychology may be the predominant providers of child custody evaluations we cannot simply address issues

in child custody evaluation to that subgroup. Whether that relates to conference offerings, journal articles, or in regulatory oversight, there is a diversified community of professionals providing child custody evaluation, and to approach such evaluations from other standpoints is to have blinders in place. This is a particularly important issue for educational institutions to be aware of as well. If the institutions we trust to train up the next generation of professionals perpetuate outmoded perceptions of practice, focusing on professional monoculture rather than interdisciplinary collaboration in ways that would better serve providers and families alike, they risk becoming irrelevant to sub-specialty fields like child custody evaluation.

By and large evaluators reported they were of Caucasian race/ethnicity. It seems unlikely this paucity of non-Caucasian evaluators is consistent with the broader communities in which most evaluators practice. Communities, both professional and otherwise, benefit from diversity of opinion and experience, diversity which may be missing when a single group is dominant. Although specific outreach strategies are beyond the scope of this work, it appears clear that attempts to broaden the pool of evaluators to better reflect the communities they serve may be warranted was well.

Finally, there appears to be clear cost savings involved in utilizing masters-level evaluators that will need to be considered by policy makers of all levels. If masters-level evaluators can more efficiently provide the same services as more expensive doctoral-level practitioners it seems clear, in an environment where funds available to courts, other government units, and individual families are limited, that recruitment and training should be directly focused towards that degree tier. If there are cases that are better suited towards evaluators of different educational backgrounds then both individual and systematic explorations of goodness of fit

between evaluator and custody-disputing family must be made. It seems likely that in families lacking in significant pathology the advanced psychometric training of many doctoral-level providers may be an unnecessary premium to pay for. Likewise for circumstances where the issues the family is experiencing are related primarily to dysfunction in their family system (which have lead to separation or a disrupted transition of in their family system to a healthy two household co-parenting arrangement), a masters-level evaluator may be an appropriate selection. Even in cases where there is known serious mental illness, substance abuse, or other behavioral concerns if those are secondary concerns to the court in relation to the family transition issues that are being litigated and examination of the individual evaluator's training and resulting assignment (or not) to a case should be considered.

Research

The next stage in research on practice issues seems to be to start asking about more than simple historical specialty training ("have you ever" type questions) to also investigating evaluator's ongoing training – how they stay abreast of a field that evolves both in terms of evolving understandings of human and societal issues, but also one where the very laws evaluators are asked to frame their work product within are changing. There also appear to be specific training issues, such as domestic violence, sexual violence, and child abuse and neglect that should be regularly assessed in additional research projects. Particularly of interest may be a qualitative review of evaluator practices and training in order to explore in a rich and contextual manner their viewpoints. While such studies may be seen as prohibitively expensive to do on large scales, it may be possible to examine best practices through targeted interviewing of

selected experts or acknowledged “master evaluators” in the way other studies have examined insights of “master” therapists.

As noted previously, there appears to be a need to assess quality of services from various perspectives. One clear way to approach this is via direct reviews of evaluator reports, integrating the information obtained into current and evolving research. As noted previously, the two most recent North American examples of such work, Bow & Quinnell (2002) and Horvath, Logan & Walker (2002), in addition to being dated, employed divergent methodology and studied seemingly disparate samples of evaluators. Clearly breaking down not only work product reviews but future survey research (and perhaps eventually reanalysis of this study’s data set) by employment arrangements (working in a court agency versus providing services in a private practice), in addition to educational level will allow a more apples-to-apples comparison of evaluators. It may also help to explicate if there are differences in which factors are valued (and how) for those in situations where the performance demands may also be subtly different.

Finally there appears to be a clear high priority area to target for future research: isolating differences in costs between various private practitioners of different educational and theoretical backgrounds. While many families of limited means may be assigned to particular Court Services staff member for evaluation, where there is the possibility of choice (or where there is a lack of social supports such as Court Services offices leading only to evaluation by private practitioners) a consumer (or more likely, their attorney) who understand the utility of various procedures may make a more informed decision on where resources are committed. This may also aid evaluators, who are presumably offering the highest quality services they can manage, to make decisions about which procedures are optimal for obtaining data necessary for making

recommendations to the courts. This might include issues similar to a dose response paradigm, where efficacy or increases to incremental validity may peak or taper off at particular points. On the other hand further research into this area may show that an adaptive approach, fitting each family's circumstance, has the most utility. As Tolstoy observed in his masterpiece Anna Karenina, "Happy families are all alike; every unhappy family is unhappy in its own way."

Appendix A

Tables summarizing previous studies

Table A-1 Comparative breakdown of study purpose, research method, sample method, sample size, and analysis

Study	Purpose/Applicability to Custody Evaluation	Research Method	Sample Method	Response Rate / Sample Size	Comparison Group(s)	Statistical Analysis
Abrams (1988)	Examined clinician response patterns in contrast to clinician's own personal divorce history. (Clinician = 50% of time in clinical services.)	Survey	Availability [geographic] (nonprobability)	900 questionnaires sent, 210 returned (23% rate), 123 usable returned. All respondents clinical psychologists	None (in-group 3x3 factoral design)	MANOVA, ANOVA
Ackerman & Ackerman (1997)	Replication of earlier study (Keilin & Bloom, 1986) with additional question items	Survey	Purposive (nonprobability)	800 questionnaires mailed, 338 returned (42% rate), 201 returned fit final selection criteria (25% rate) All respondents doctoral level psychologists	None	M and SD of endorsed items
Ackerman & Brey Pritzl (2011)	Replication of earlier studies (Ackerman & Ackerman, 1997; Keilin & Bloom, 1986) with additional question items	Survey	Purposive (nonprobability)	213 questionnaires completed[a] All respondents doctoral level psychologists	None	Majority is simple listing of responses

[a] Although the published article notes these were "filled out completely" this conflicts with the underlying dissertation which the article is based on, which notes 213 responses used, although not all 213 were fully completed. Additionally, although the study lists this as 26.7% of 800 "requests made" at least one source of respondents came from a broadcast e-mailing to a professional organization. It is difficult to tell with what amounts to a snowball sampling method how a response rate was calculated and it is therefore omitted here.

Ackerman & Steffen (2001)	Explored judges' expectations of the practices from Ackerman's 1997 study	Survey	Purposive (nonprobability)	800 sent, response rate unclear All respondents family judges	None	Majority is simple listing of responses, M and SD of sole/joint custody factors
Austin, Jaffe, & Friedman (1994)	Compared levels of experience and background characteristics to the number of strategies used to compensate for evaluator bias	Survey	Purposive (nonprobability)	800 questionnaires mailed, 174 useable returned (21.75%) 81 Social workers 78 Psychologists 14 Psychiatrists 1 Multiple licenses?	Yes - Clinicians with no evaluation experience	ANOVA, chi-square, post hoc Tukey test, f-test, t-test
Bow & Boxer (2003)	Examine training and practices in relation to domestic violence issues in evaluations	Survey	Availability (nonprobability)	348 potential, 115 responded 18 Social workers 18 Masters level psychology/counseling 78 doctoral level psychologists	None	M and SD of endorsed items
Bow & Quinnell (2001)	Evaluate level of adherence to APA guidelines for child custody evaluation, compared results to Ackerman & Ackerman (1997), Keilin & Bloom (1986), and LaFortune & Carpenter (1998)	Survey	Purposive (nonprobability)	563 questionnaires mailed, 279 returned (50%), 198 fit selection criteria (35%) Masters or doctoral level psychologists only, 96% doctoral level	None	M and SD of endorsed items

Bow & Quinnell (2002)	Review 52 actual reports from across the USA for congruence to survey data, quality, and how communicated to the court	Record review	Availability (nonprobability)	265 initially contacted, 78 interested, only 56 responded, 52 met criteria All reports by doctoral level psychologists	None	Simple listing of responses
Bow & Quinnell (2004)	Examined congruency between evaluations and needs of legal profession, both 1. attorneys and 2. judges	Survey	1. Random 2. Not a sample, all family judges in State of Michigan contacted.	1. 300 Sent, 120 returned (40%), 89 usable. 2. 124 sent, 37 returned (30%) All judges and attorneys	Compared group 1 and group 2 responses	MANOVA, M and SD of endorsed items
Bow, Quinnell, Zaroff, & Assemany (2002)	Evaluate level of adherence to child sexual abuse guidelines when evaluating alleged victims in SAPCR	Survey	Purposive (nonprobability)	368 questionnaires mailed, 147 returned (40%), 84 returned fit selection criteria (22.8%) 98% doctoral level psychologists 2% Unknown	None	M and SD of endorsed items
Brandt, Dawes, Africa & Schwartz (2004)	Review of actual reports to examine substantive issues informing evaluator decision making	Thematic content analysis	Convenience (nonprobability)	39 reports from 8 psychologists (unknown sub-types)	None	Thematic content analysis

Caplan & Wilson (1990)	Early survey of practitioners. Canadian setting. Looked at differences between disciplines. "Forgotten" study – seldom mentioned.	Survey	Geographic, not sample w/in geography (survey sent to each licensed person)	4,700 questionnaires sent, 220 usable (met criteria) returned 45% social workers 34% psychologists 18% psychiatrists	None	Frequency analysis, Chi-square & qualitative
Cohen & Shnit (2001)	Explore social workers' parenting time recommendations for non-custodial fathers	Record review	Purposive (nonprobability)	151 reports (all reports that met the sample criteria for 1993 & 1994) All by social workers	None	M, % & chi-square on coded items, stepwise multiple regression
Crosby-Currie (1996)	Examine children's involvement in custody cases based on 1. judicial, 2. attorney and 3. mental health professionals experience	Survey	1. Not a sample, all family judges in states studied contacted. 2. Random 3. Random	1. 281 sent, 37.4% reply 2. 414 sent, 38.4% reply 3. 533 sent, 28.7% reply MHPs were psychologists only	Compared groups 1, 2 & 3 with each other	MANOVA, chi-square
Davidson-Arad, Cohen & Wozner (2003)	Explored basis for custody recommendations as a function of perceived quality of life little impacted by parental features/characteristics; low SES & pathology more concerning with fathers. Review of actual CCEs.	Survey	Convenience (non-probability)	130 Social workers,57% BA, 34% MSW, 9% PhD All by social workers	Within group comparison by gender	Chi square, ANOVA

Fitzgerald & Moltzen (2004)	Report of survey data from psychologists in New Zeeland performing CCEs	Survey	Unclear how sample chosen	111 surveys sent, 31.5% returned Psychologists only	None	Simple frequency analysis; M & SD of vignette responses
Gourley & Stolberg (2000)	1. Assess how many psychologists conduct custody evaluations 2. Survey 'credible' evaluators for rates of agreement on importance of issues	Survey	1. Random, 2. Purposive (nonprobability) *Note: When asked for "credible" psychologist evaluators 43.7% of those nominated were not actually psychologists*	1. n = 186 (implies response rate of 100%) 2. 49 questionnaires sent, 27 responses received (55%) of which 21 were useable (42.8%) Psychologists only	Compared responses from group 1 with group 2	M and SD of endorsed items
Hagen & Castagna (2001)	Reanalysis of Ackerman & Ackerman (1997) figures regarding test frequency indicating much lower rates of use when testing frequency looked at.	Re-analysis	NA, but important to field in correcting data misimpressions A&A 1997 gave	Same as Ackerman & Ackerman (1997) All respondents doctoral level psychologists	--	Group frequency analysis with better rigor

Horvath, Logan, & Walker (2002)	Assess adherence to general CCE guidelines and examine differences based on evaluator training	Record review	Random	n = 82 64.7% Friend Of Court (social workers) 31.4% private evaluators (mixed 66% Ph.D. psychologist, 25% MSW, 9% Ed.D. psychologist) 3.9% CPS (excluded due to small number)	Group analyzed as a whole and then as two groups divided by private evaluators and court employees	Content analysis - M and SD of observed items, chi-square
Hughes & O'Neal (1983)	Short article. Much flawed methodology/assumptions. Not functional enough to be used meaningfully other than as a passing mention.	Survey	Convenience	Surveys sent through inpatient treatment centers, authors acknowledge this was a failing on their part to account for diversified practice areas.	None	Simple report of percentages
Jaffe & Cameron (1984)	Early attempt at examining methods chosen for resolving custody cases, including evaluation Not all that useful in current context due to changing social factors.	Survey	Availability (nonprobability)	n = 40 cases Mixed involvement, unclear what percentage of cases from which discipline.	Sample divided by intervention type (assessment, arbitration, mediation)	M and SD of endorsed items, chi-square

Jameson, Ehrenberg & Hunter (1997)	Assessed importance of different factors to concept of "best interest of child"	Survey	Geographic, surveys sent to all psychologists listing a divorce or custody specialty in British Columbia registry	88 surveys sent, 78 returned All psychologists	None	Principal-components analysis
Keilin & Bloom (1986)	Examine basic practices, factors viewed as important in decision making.	Survey	Purposive (nonprobability)	302 questionnaires mailed, 190 returned (62.9%), 82 fit selection criteria (27%) 78% Doctoral level psychologists, 18% psychiatrists, 2% "masters level practitioners" and 1% social workers	None	M and SD of endorsed items
Kunin, Ebbesen & Konecni (1992)	Examined factors in reports that impacted judicial decision making. Sample is from 1982, pre-dates many major changes in law and practice.	Record review	Geographic (all available reports in San Diego County)	282 reports Unclear training of "court counselors"	None	Log-linear analysis, chi-square

LaFortune & Carpenter (1998)	Assess practices and attitudes of mental health professionals in custody evaluations, adherence to guidelines, and replicate earlier study (Keilin & Bloom, 1986) with additional question items	Survey	Purposive (nonprobability)	286 questionnaires mailed, 165 completed (61%) 89% Psychologist (87% doctoral level, 2% masters) 7% LPC/MFT 3% social work	None	M and SD of endorsed items, correlation
Logan, Walker, Jordan & Horvath (2002)	Compared CCE procedures in DV and non-DV cases	Record review	Random	N=82 Appears to be same dataset as Horvath et al, but this is not explicated	DV (56%) vs non-DV (44%)	Chi-square
Lowery (1985)	Explored importance of various criteria to mental health professionals in CCE cases; examined experienced vs. inexperienced professionals ratings	Survey	Random	100 psychologists & 100 SW sent surveys 50 psychologists (55% doctoral level) & 54 social workers responded (0% doctoral level)	None	Mean & SD of endorsed items
Quinnell & Bow (2001)	Same data set as Bow & Quinnell (2001) looking at psychometric testing. Compared results to Ackerman & Ackerman (1997), Keilin & Bloom (1986), and LaFortune & Carpenter (1998)	Survey	Purposive (nonprobability)	563 questionnaires mailed, 279 returned (50%), 198 fit selection criteria (35%) Masters or doctoral level psychologists only, 96% doctoral level	None	M and frequency of endorsed items

Sagi & Dvir (1993)	Assess value-bias of female custody evaluators	Survey	Purposive (nonprobability)	n = 216 All social workers	Participants randomly assigned to one of four theoretical cases	ANOVA, chi-square, Tukey test, Friedman test, Nemenyi test
Schindler (1985)	Examined court orders for detail Information now very outdated due to societal changes	Record review	Unclear	n=62 All cases involving social workers	None	Descriptive only

Table A-2 – Custody evaluator procedures

Procedure	Mean hours spent in procedure (% of respondents using this procedure, where reported)					
	Keilin & Bloom (1986) [a]	Ackerman & Ackerman (1997)	LaFortune & Carpenter (1998)	Gourley & Stolberg (2000)[b]	Bow & Quinnell (2001)	Ackerman & Brey Pritzl (2011)
Interviewing parents	4.07 (100%)	4.7	4.85	5.3	7.0+[c]	7.1
Interviewing children	1.55 (98.8)	2.7	1.44-1.89[d]	3.5	1.75	3.6
Psychometric testing of adults	2.66 (75.6%)	--	--	--	3.03 (91%)	--
Psychometric testing of children	2.50 (74.4%)	--	--	--	1.97 (61%)	--
--Combined adult and child testing	--	5.2	--[e]	4.0	--	6.1
Parent-child observation	1.89 (68.8%)	2.6	2.49	3.0	1.59 (92%)	3.7
Interviewing significant others[f]	1.32 (48.8%)	1.6	1.37	2.1	--	2.3
Collateral contacts[g]	--	1.6	1.16	2.2	--[h]	3.2

[a] Respondents who reported not using a particular procedure were not factored in to this calculation.
[b] Data on procedures was only reported from the sample of evaluators rated as "credible" rather than a general sample of evaluators
[c] Like Keilin & Bloom this study examined the percentage of respondents using particular procedures, however it is unclear how they aggregated this information to get the "+" (apparently meaning 7 hours and more) in this figure.
[d] Dependent on age
[e] LaFortune & Carpenter present data on testing on a test by test basis, but it is unclear from their figures the mean hours for all testing procedures
[f] "Significant others" is not adequately operationalized in several of the studies – in Keilin & Bloom (1986), for instance, it is parenthetically noted that this means "friends and relatives" (p.340) whereas no such clarification is provided in later studies. It is unclear how this differs in some studies from "collateral contacts" although it appears to mean spouses and paramours in later studies.

Home visits	1.42 (30.0%)	--	--	--	2.14 (33%)	--
Reviewing materials	--	2.6	1.11[i]	2.4	2.97 (98%)	5.6
Report writing	2.84 (93.8%)	5.3	--	4.3	7.3	10.6
Consulting with attorneys	1.36 (91.4%)	1.2	--	1.9	1.31 (28%)	1.3
Court testimony	2.29 (79.0%)	2.2	--	2.5	--	2.6

Dashes indicate this item was not assessed or if assessed not reported.

[g] This term is not consistently defined, although various studies focused on interviews and/or questionnaires used with third parties who may have relevant information. It is generally used to mean friends, non-nuclear family members, or professional contacts in line with prevailing nomenclature. (Austin, 2002).

[h] Bow & Quinnell break down collateral contact by several sub-groups: therapists, teachers, physicians, neighbors and friends, and relatives but do not provide a clear cumulative figure.

[i] Limited to school records

Table A-3 Evaluator demographics and experience

Respondent variable	Study: Keilin & Bloom (1986)[a]	Ackerman & Ackerman (1997)[b]	LaFortune & Carpenter (1998)	Gourley & Stolberg (2000)[c]	Gourley & Stolberg (2000)[d]	Bow & Quinnell (2001)[e]	Ackerman & Brey Pritzl (2011)[f]
Number of respondents	82	201	165	65	21	198	213[g]
Average age (years)	47.7	49.1	--	--	--	51 years	56.2
Gender: Male / Female	64 / 18	139 / 62	97 / 68	--	17 / 4	103 / 95	130/83
Average number of evaluations completed (lifetime)	156.5	214.9	98.2	--	--	245	269
Average number of evaluations in last year	16.9	--	8.1	2.47	--	--	--
Education:							
Doctoral level	96.4%	100%	92.8%	--	--	96%	100%
Masters level	3.6%	0%	7.2%	--	--	4%	0%
Profession:							
Psychologist	78.1%	100%	89%	100%	100%	100%	100%
Social worker	1.2%	0%	3%	0%	0%	0%	0%
Professional counselor	2.4%[h]	0%	6%	0%	0%	0%	0%
Marriage & family therapist	--	0%	1%	0%	0%	0%	0%
Psychiatrist	18.3%	0%	0%	0%	0%	0%	0%
Respondent variable	Keilin & Bloom	Ackerman & Ackerman	LaFortune & Carpenter	Gourley & Stolberg	Gourley & Stolberg	Bow & Quinnell	Ackerman & Pritzl (2011)[n]

[a] Inclusion criteria required involvement in at least 4 custody evaluations
[b] Inclusion criteria required involvement in at least 10 custody evaluations
[c] Using a general sample
[d] Using a sample of evaluators rated as "credible" rather than a general sample of evaluators
[e] This sample is identical to the companion paper by Quinnell & Bow (2001).
[f] Minimum number of evaluations completed to be included was 10 custody evaluations
[g] Although the published article notes these were "filled out completely" this conflicts with the underlying dissertation which the article is based on, which notes 213 responses used, although not all 213 were fully completed.
[h] Listed as "masters level practitioners" in the study

	(1986)[i]	(1997)[j]	(1998)	(2000)[k]	(2000)[l]	(2001)[m]	
Experience:							
Average years in practice	16.1	19.0	17.9	--	--	22.66	24.8
Years in general forensics	--	--	--	--	--	15.62	--
Years in child custody field	10.6	--	9.6	9.8	--	13.57	--
Primary training with both adults and children	--	--	87.3%	--	--	76%	--
Training in child custody derived from:							
--Seminars	--	--	--	43.0%	57.1%	86%	95%
--Supervision	--	--	72.1%	36.9%	28.6%	44%	--
--Internship or postdoc	--	--	30.3%	12.3%	14.3%	39%	13.3%
--Graduate forensic course	--	--	21.2%	13.8%	33%	18%	12.8%
--Reading books/journals on custody evaluation	--	--	--	78.5%	76.2%	--	--
--Research in the field	--	--	--	[o]	23.8%	--	--
--Other training	--	--	--	21.5%	--	16%	4.3%[p]

Dashes indicate this item was not assessed or if assessed not reported. Studies indicating 100% of respondents were psychologists surveyed only psychologists as part of their selection criteria.

[n] Minimum number of evaluations completed to be included was 10 custody evaluations
[i] Inclusion criteria required involvement in at least 4 custody evaluations
[j] Inclusion criteria required involvement in at least 10 custody evaluations
[k] Using a general sample
[l] Using a sample of evaluators rated as "credible" rather than a general sample of evaluators
[m] This sample is identical to the companion paper by Quinnell & Bow (2001).
[o] The study combined this data into the "other" field along with "clinical" experience and other undisclosed factors. It is unclear why research was separated out in the "credible" sample but no comparison with the general sample was given.
[p] Over 300 practicum hours

Table A-4 Caplan & Wilson (1990) information in tabular format

Sub-table A-4A – Potential Sources of Bias and Conflicts of Interests

Issue	Percentage of respondents endorsing item
My own sex [gender] is relevant to my assessment approach	30
My professional discipline is relevant	82
Whether or not I have children is relevant	49
My age is relevant	37
Whether or not I have been abused or neglected is relevant	39
Whether my parents had been separated or divorced during my childhood is relevant	33
Whether I am foreign-born, whether English is my first language, and whether I am a member of a minority group are relevant	39.5
My sexual orientation is relevant	21
It affects my assessment if I know one of the lawyers socially or have worked with the spouse of one of the lawyers	15
I take care never to talk to one lawyer without the other being present	14.5
The reputation of, or my knowledge about, either lawyer affects my assessment	20.5

Sub-table A-4B – Beliefs and Principles

Issue "All other things being equal…"	**Percentage of respondents endorsing item**
…infants and young children should live with their mothers	"nearly half"[*]
…male children should live with their fathers	7
…female children should live with their mothers	19
…a parent who was abused as a child should not be the residential parent	22
…the parent who can provide greater financial security should be the residential parent	19
…the residential parent's sexual orientation should not be homosexual	25.5
…the parent who can minimize inter-parental conflict should be the residential parent	88
…the parent who can provide a two-parent home should be the residential parent	14
…the more fragile or disturbed parent should be the residential parent	2
…the parent to whom the child has less of an attachment should be the residential parent	55
…preadolescents rarely lie in situations of alleged parental sexual abuse of children	60.5
…adults of either sex rarely lie when they say their ex-spouse has sexually assaulted or hit them	"less than one third"

[*] Caplan and Wilson are frustratingly imprecise at times.

Sub-table A-4C – Practices

Issue The evaluator	**Percentage of respondents endorsing item**
regularly does not see children who are the subject of the dispute	8
regularly sees each parent with the child or children in question	80
sees parents' new spouses or partners with the children	73
regularly obtains third-party information	75
regularly obtains information from previously involved clinicians	71
does not regularly ask if the child in question was planned or unplanned	25
does not ask if the child was wanted or unwanted	29.5
does not inquire about whether the parents have criminal histories	18
does not inquire about whether parents have psychiatric histories	13
does not request to see relevant court documents	43
believes it important to have some understanding of crucial legal issues related to child custody	77

Sub-table A-4D – Proportion of evaluators who have <u>not</u> read relevant clinical literature in the last two years regarding various practice areas

Practice Area	**Percentage of respondents endorsing item**
Whether abused children are likely to become abusive parents	86
Whether mothers at home tend to provide better nurturing than working mothers	50
Whether preadolescent children whouls be exposed to the fact that they have a homosexual parent	29
Whether an infant child requires continuity with a single "primary" parent	67
Whether children who are alleged victims of abuse rarely lie	72

Table A-5 Psychometric testing percentages (percentage of evaluations for which a specific test was used)

Hagen & Castagna (2001) re-evaluated the Ackerman & Ackerman (1997) data to obtain the actual percentage of the more than 43,000 evaluations that made up their dataset (201 evaluators with 214.9 average evaluations per evaluator) in which a particular psychometric test was used. They point out "no test other than the MMPI was used in even one third of these evaluations" (pg. 271) and that the format the numbers were previously presented in represented a misleadingly high estimation of the use of such tests.

A similar reanalysis of selected items from Ackerman & Ackerman (1997) along with Keilin & Bloom (1986) and Quinnell & Bow (2001) is presented below.

Ackerman & Brey Pritzl (2011) did not report frequency of use data, and thus a comprehensive analysis of their information is not possible.

	# of evaluators / # of cases / % of total evaluations								
Test	Keilin & Bloom (1986)[a]			Ackerman & Ackerman (1997)			Quinnell & Bow (2001)		
Intelligence Tests									
Wechler Adult Intelligence Scale-R/III	24	105	20%	86	105	21%	93	76	15%
Child IQ test[b]	37	133	38%	117	97	26%	95	78	15%
Academic Tests									
Wide Range Achievement Tests-R/3-Adult	--	--	--	20	168	8%	20	103	4%
Child achievement test[c]	17	119	16%	56	120	16%	51	91	10%
Objective personality tests – adults									
MMPI-1/2	58	137	62%	185	196	84%	186	216	83%
MCMI	--	--	--	68	157	25%	103	179	38%
16 Personality Factor	5	94	4%	16	144	5%	178	115	42%
Objective personality tests - adolescents									
MMPI-A	6	76	4%	40	105	10%	85	103	18%
Millon Adolescent Clinical Inventory	--	--	--	22	88	4%	42	88	8%

[a] Includes only those respondents who reported using the test in mean % of cases calculation
[b] Child intelligence tests include the Wechler Intelligence Scale for Children (3rd ed.), Stanford-Binet (4th ed.) Kaufman ABC, and McCarthy Scales of Cognitive Abilities
[c] Child's achievement tests include Wide Range Achievement Test-R/3, Wechsler Individual Achievement Test, Peabody Individual Achievement Test, and Kaufman Test of Educational Achievement

Appendix B

Sample outreach letters

Outreach Letter to an Organization/Entity

<Date>

<Organizational contact name>
<Contact address>

Dear <Contact name>,

We writing in regards to a research project on child custody evaluators that is being completed as part of doctoral studies at the University of Texas at Arlington School of Social Work. In order to complete the research, a survey of professionals who perform child custody evaluations is needed.

We are seeking information on how to contact members of <Organization> to participate in the study. If you can provide us with the information needed to obtain a list of your members who perform child custody evaluations, access to any e-mailing lists or discussion groups you run where information about the study might be posted, or any other service in disseminating this information to those who might be able to participate it would be greatly appreciated.

Thank you for your time in this matter. Please follow up with us in whatever manner is most convenient to you:

Telephone: 817-239-3828

Fax: 940-343-2601

E-mail: aaron@texascounseling.org

Sincerely,

Aaron Robb
University of Texas at Arlington
School of Social Work Ph.D. candidate

Maria Scannapieco, Ph.D.
Dissertation committee chair
<contact info>

Outreach Letter to an Individual

Dear professional,

Your help is requested in a research project regarding the practices of professionals of all backgrounds who conduct child custody evaluations (also known as parenting time evaluations, child custody studies, and other names in various jurisdictions). We recognize as a forensic professional you are likely already quite busy, and have designed the research questionnaire to be as easy to complete as possible. It should only take approximately 20 minutes to complete, and will provide important data regarding child custody evaluation.

The questionnaire is designed to be filled out anonymously. It is administered online through the web site "survey monkey." You do not need to print or mail anything, simply complete it on line and when done click on the submit button. We are trying to reach as many child custody evaluators as possible with this research. Please feel free to pass this information along to your colleagues and disseminate this request for participation as widely as possible. We are asking for your help in casting this "broad net" as there is no clear consensus as to how many child custody evaluators there actually are, let alone how to reach them. Each additional evaluator who contributes helps make the research results more meaningful through the data they provide.

The research results will be used in the completion of a dissertation project and then submitted for formal publication to peer-reviewed journals. If you have any questions regarding this study, or if you wish to have a copy of the results of the study, please e-mail Aaron Robb at aaron@texascounseling.org. Alternatively you may contact Mr. Robb via telephone at 817-239-3828, or reach Dr. Maria Scannapieco at <contact info> or <human subjects research line here>.

Thank you in advance for your participation.

Sincerely,

Aaron Robb
University of Texas at Arlington
School of Social Work Ph.D. candidate

Maria Scannapieco, Ph.D.
Dissertation committee chair
<contact info>

The questionnaire can be reached via: <link>

We are keeping the survey open for a limited time, please submit your responses no later than <date>

Appendix C

Survey Questions

Jurisdictional language note: Thank you for agreeing to participate in this survey. In case there is any confusion regarding jurisdictional definitions of "child custody evaluation" we are referring, collectively, to evaluations by a neutral mental health professional regarding allocation of post-separation parenting arrangements as child custody evaluations. If you see other terms which have both broad common meanings and more precise or idiosyncratic legal meanings in your jurisdiction please respond based on the broader meaning if possible.

I. Demographic Data

Age: _____ Gender: _____ Race: _____

Location type (mark one): Urban / Suburban / Rural

What state(s) do you practice in: _____

Highest degree: _____ Field of study: _____

Please list any professional associations to which you belong: _____

_____ Years of experience in mental health

_____ Years of experience in forensics

_____ Years of experience in child custody

_____ Percentage of your practice devoted to child custody evaluations

_____ Total number of child custody evaluations completed in your career

_____ Number of custody evaluations completed in the last 12 months

What type of training in custody evaluations have you received? Mark all that apply

A. Graduate course(s) in forensic mental health

B. Graduate course(s) in assessment of children & families

C. Practicum

D. Internship/post-graduate placement experience, 100% custody evaluation

E. Internship/post-graduate placement experience, partial custody evaluation

F. Books/journal articles

G. Seminars and workshops

H. Post-graduate supervision

I. Research in child custody

J. Other: _____

Forensic setting:

____ Private practice ____ Court clinic/agency ____ Public mental health clinic

____ University clinic ____ Other

Have you ever participated in a survey regarding your child custody evaluation practices before?

Y/N

II. Custody Evaluation Practices

A. Employment capacity

1. I prefer to be (mark one)

a. retained by one parent/attorney

b. retained by both parents/attorneys

c. appointed by the court or guardian ad litem

d. Other_____

2. What is the percentage of child custody evaluations you perform that are court ordered? _____

B. Custody evaluation areas

Please indicate the average/typical amount of total time spent in the following areas for a complete child custody evaluation:

AREA	TIME SPENT (IN HOURS)
Interviews	
--Interviewing parents	_____
--Interviewing children	_____
--Interviewing significant others	_____
--Other interviews (outside of collateral contacts)	_____
Collateral contacts	
-- Collateral contacts with teachers/schools	_____
-- Collateral contacts with neighbors or friends	_____
-- Collateral contacts with family physicians or pediatricians	_____
-- Collateral contacts with relatives	_____
-- Collateral contacts with therapists and psychiatrists	_____
-- Collateral contacts with law enforcement related agencies	_____
-- Collateral contacts with others	_____
Parent – child observations in office/playroom	_____

Parent – child observations during home visits _____

Home visits (other than parent-child observation time) _____

Reviewing materials _____

Collateral contacts _____

Psychometric testing of adults _____

Psychometric testing of children _____

Report writing _____

Meeting with parents to review findings/recommendations _____

Meeting with attorneys to review findings/recommendations _____

Consulting with attorneys (other than to review findings/recommendations) _____

Testifying in court _____

C. Behavioral observations

1. Format – Which of the following approaches do you typically use in your child custody evaluations (mark all that apply):

a. Observe each parent with each child alone

b. Observe each parent with all children together

c. Observe both parents with each child alone

d. Observe both parents with all children together

e. Observe the children with stepparents

f. Observe the children with a parent's live-in partner?

g. Observe the children with significant others not residing in the home

2. Type of activities during observation (mark one)

___ Unstructured

___ Structured task provided

___ Combination of unstructured and structured

3. Location of observation

____ In child's homes

____ In evaluator's office

____ Both in home and in office

____ Other: _____

D. Collateral contacts

1. Do you typically obtain information from professional collateral sources (schools, physicians, therapists, etc.) via

a. Written communication

b. Interviews

c. Both written information and interviews

d. I do not obtain information from these collateral sources

2. Do you typically obtain information from non-professional collateral sources (family friends, relatives, etc.) via

a. Written communication

b. Interviews

c. Both written information and interviews

d. I do not obtain information from these collateral sources

E. Psychometric testing with children/adolescents

Please indicate the percentage of time that you currently use each of these tests in custody evaluations (i.e. If you currently use the Wechsler Intelligence Scale for Children-4th Edition, but only administer it to 70% of your child custody evaluation cases, then you would list note 70% for WISC-IV). If you do not currently use the test please mark 0%.

Please indicate the percentage of time you use the psychometric instrument as a test. Please do not include use of any instruments designed to be rated or scored that you use as interviewing tools rather than as formal tests.

General Tests	% of time used
Achenbach Child Behavior Checklist (CBCL)	_____
Behavior Assessment for Children (BASC)	_____
Bricklin Perceptual Scales	_____
Bender-Gestalt	_____
Children's Apperception Test/Thematic Apperception Test (CAT/TAT)	_____
Children's Depression Inventory (CDI)	_____
Conners Rating Scale	_____
Family Relations Test	_____
Millon Adolescent Personality Inventory (MAPI)	_____
Minnesota Multiphasic Personality Inventory – Adolescent (MMPI-A)	_____
Perceptions of Relationships Test (PORT)	_____
Personality Inventory for Children (PIC)	_____

Projective Drawings _____

Roberts Apperception Test _____

Rorschach _____

Sentence Completion _____

Other_____

_____

_____

Cognitive Tests % of time used

Achievement Test (Wechsler Individual Achievement Test II (WIAT II);

Woodcock-Johnson Tests of Achievement – Third Edition;

Wide Range Achievement Test 3 or 4 (WRAT3/WRAT4)) _____

Intelligence Test (Wechsler Intelligence Scale for Children IV (WISC IV);

Stanford Binet 5th Edition; Kaufman Assessment Battery for Children-2

(KABC-2); Wechsler Abbreviated Scale of Intelligence (WASI)) _____

Paper and Pencil IQ Test _____

Other_____

_____

_____

F. Psychometric testing with adults

Please indicate the percentage of time that you currently use each of these tests in custody evaluations (i.e. If you currently use the Wechsler Intelligence Scale for Children-4th Edition, but

only administer it to 70% of your child custody evaluation cases, then you would list note 70% for WISC-IV). If you do not currently use the test please mark 0%.

Please indicate the percentage of time you use the psychometric instrument as a test. Please do not include use of any instruments designed to be rated or scored that you use as interviewing tools rather than as formal tests.

General Tests	% of time used
16 Personality Factor (16 PF)	_____
Beck Depression Inventory	_____
Bender-Gesalt	_____
Michigan Alcohol Screen Test	_____
Minnesota Multiphasic Personality Inventory 2 (MMPI-2)	_____
Minnesota Multiphasic Personality Inventory 2 Restructured Form (MMPI-2-RF)	_____
Millon Clinical Multiaxial Inventory III (MCMI-III)	_____
Personality Assessment Inventory (PAI)	_____
Projective drawings	_____
Rorschach	_____
Sentence Completion	_____
Thematic Apperception Test	_____

Other_____

_____

_____

Cognitive Tests	% of time used
Achievement Test (Wechsler Individual Achievement Test II (WIAT II); Woodcock-Johnson Tests of Achievement – Third Edition; Wide Range Achievement Test 3 or 4 (WRAT3/WRAT4))	_____
Intelligence Test (Wechsler Adult Intelligence Scale – III (WAIS-III); Stanford Binet 5th Edition; Wechsler Abbreviated Scale of Intelligence (WASI))	_____
Paper and Pencil IQ Test	_____
Wechsler Memory Scale-III	_____
Other_____	

_____

_____

Parenting Tests	% of time used
Ackerman – Schoendorf Scales for Parent Evaluation of Custody (ASPECT)	_____
Adult Adolescent Parenting Inventory	_____
Child Abuse Potential (CAR)	_____
Custody Quotient	_____
Parent – Child Relationship Inventory (PCRI)	_____
Parenting Stress Index (PSI)	_____

Parent Awareness Skills Survey (PASS) _____

Other_____

_____

_____

G. The percentage of time that you use co-evaluators: _____

Preference when using a co-evaluator (mark your preferred choice)

1. Same sex co-evaluator

2. Opposite sex co-evaluator

3. No preference

H. Fees and practice management issues

1. How do you charge in a typical case

___ Per hour? If so, how much $ _____

___ Per case? If so, how much $_____

2. Do you have a separate fee for psychometric testing?

___ Per hour? If so, how much $ _____

___ Per case? If so, how much $_____

___ I do not have a separate fee for psychometric testing

3. What percentage of the fee do you require by:

____ the first appointment

____ the mid-point of evaluation

____ by the final appointment

____ by the time the report is completed

(Total should = 100)

4. In what percentage of cases do you

___ provide a written report

___ provide a verbal report

___ provide both a verbal and written report

5. Do you provide a report if you have not received your complete fee? Y/N

6. Your average fee for a custody evaluation: _____

7. From the first interview to completion of your report how many weeks does it take to complete a child custody evaluation: _____

8. Percentage of cases you testify in court regarding child custody work: _____

9. How do you charge for testimony?

___ Per hour? If so, how much $ _____

___ Per case? If so, how much $_____

___ I do not have an additional fee for testimony

10. Payment arrangements for testimony

a. Full payment received prior to testimony

b. Partial payment received prior to testimony

c. No requirement

11. % of cases you give depositions in court regarding child custody work: _____

12. How do you charge for depositions?

___ Per hour? If so, how much $ _____

___ Per case? If so, how much $_____

___ I do not have an additional fee for depositions

13. Payment arrangements for deposition

a. Full payment received prior to deposition

b. Partial payment received prior to deposition

c. No requirement

_____Number of licensing board complaints have been filed against you regarding your custody work?

____ Number of times a malpractice suit has been filed against you regarding your custody work?

I. Do you expect that you will:

Review pleading of family law case?	Do not expect, Neutral, Expect
Review legal records?	Do not expect, Neutral, Expect
Review criminal records?	Do not expect, Neutral, Expect
Review children's school records?	Do not expect, Neutral, Expect
Review parents' medical records?	Do not expect, Neutral, Expect
Perform psychometric testing on a parent's significant other?	Do not expect, Neutral, Expect
Perform a home visit?	Do not expect, Neutral, Expect

Consult with a guardian ad litem?	Do not expect, Neutral, Expect
Consult with parties' attorneys?	Do not expect, Neutral, Expect

III. Custody Decision Making

A. Legal custody

Schepard (2004) loosely defines joint custody as "a post-divorce parenting arrangement in which parents substantially share decision-making for their child (joint legal custody) and the child spends substantially equal time at each parent's residence (joint physical custody)" (pg 35). In non-technical language it follows that sole custody is the opposite of this, where parents do not share decision-making and the child spends a substantially unequal time at each parent's residence.

In this section we will ask you to focus on issues of joint and sole legal custody. For each item listed below, please indicate how important on the 9-point Likert-type scale you would consider each variable to be in deciding to recommend joint versus sole legal custody.

1	2	3	4	5	6	7	8	9
Not at all important		Unimportant		Neither important nor unimportant		Important		Extremely important

---

1. Expressed wishes of the child, age 15.

2. Expressed wishes of the child, age 10.

3. Expressed wishes of the child, age 5.

4. The age of the parents.

5. Marital status of each parent; remarried, single, or cohabiting.

6. Whether or not one parent is involved in a homosexual relationship.

7. Number of children in the family.

8. Age of the children.

9. The geographic proximity of parental homes.

10. Economic and physical similarities or differences between parental homes.

11. Economic stability of the parent.

12. Whether or not the child is placed in day care while the parent works.

13. Differences between parental discipline styles.

14. Each parent's previous involvement in caretaking responsibilities.

15. The quality of relationship the child has with each parent.

16. Psychological stability of the parents.

17. Ability of the parents to separate their interpersonal difficulties from their parenting decisions.

18. The amount of anger and bitterness between the parents.

19. Whether the child exhibits behavior problems at home or at school.

20. Amount of flexibility in parents' work schedule.

21. Influences of extended family members (e.g., in-laws and close relatives).

22. The parents' willingness to enter joint custody arrangements.

23. Differences between parents' religious beliefs.

24. Current state law (in your state).

25. Availability of extended family members.

26. Problems with the law.

27. Problems with substance abuse.

28. Cooperation with previous court orders.

29. Intelligence of the parents.

30. Gender of the child.

31. Sexual abuse allegation has been made against one of the parents.

32. Physical abuse allegation has been made against one of the parents.

33. One of the parents exhibits better parenting skills than the other.

34. The child appears to have a closer emotional bond with one of the parents.

35. One of the parents is more involved with the child than the other.

B. Physical custody

Again keeping in mind Schepard's definition above, in this section we will ask you to focus on issues of joint and sole physical custody. For each item listed below, please indicate how important on the 9-point Likert-type scale you would consider each variable to be in deciding to recommend joint versus sole physical custody.

1	2	3	4	5	6	7	8	9
Not at all important		Unimportant		Neither important nor unimportant		Important		Extremely important

---

1. Expressed wishes of the child, age 15.

2. Expressed wishes of the child, age 10.

3. Expressed wishes of the child, age 5.

4. The age of the parents.

5. Marital status of each parent; remarried, single, or cohabiting.

6. Whether or not one parent is involved in a homosexual relationship.

7. Number of children in the family.

8. Age of the children.

9. The geographic proximity of parental homes.

10. Economic and physical similarities or differences between parental homes.

11. Economic stability of the parent.

12. Whether or not the child is placed in day care while the parent works.

13. Differences between parental discipline styles.

14. Each parent's previous involvement in caretaking responsibilities.

15. The quality of relationship the child has with each parent.

16. Psychological stability of the parents.

17. Ability of the parents to separate their interpersonal difficulties from their parenting decisions.

18. The amount of anger and bitterness between the parents.

19. Whether the child exhibits behavior problems at home or at school.

20. Amount of flexibility in parents' work schedule.

21. Influences of extended family members (e.g., in-laws and close relatives).

22. The parents' willingness to enter joint custody arrangements.

23. Differences between parents' religious beliefs.

24. Current state law (in your state).

25. Availability of extended family members.

26. Problems with the law.

27. Problems with substance abuse.

28. Cooperation with previous court orders.

29. Intelligence of the parents.

30. Gender of the child.

31. Sexual abuse allegation has been made against one of the parents.

32. Physical abuse allegation has been made against one of the parents.

33. One of the parents exhibits better parenting skills than the other.

34. The child appears to have a closer emotional bond with one of the parents.

35. One of the parents is more involved with the child than the other.

IV. Recommendations

A. “Ultimate issue” questions – Which of the following do you make explicit recommendations about in your child custody evaluations (mark all that apply)

____ Legal custody

____ Physical custody (also known as parenting time, or parenting plan schedules)

____ I do not make explicit recommendations on ultimate issue questions

B. For the following items please answer on the following 5-point Likert-type scale, where 1 = strongly disagree and 5 = strongly agree

1. A child custody evaluator should be allowed to answer “ultimate issue” questions?

2. A child custody evaluator should be allowed to state who they perceive to be the better parent

3. A child custody evaluator should focus primarily on threshold issues (i.e. is parenting adequate)

4. Parenting should not be an issue unless below a certain threshold

5. Mental health should not be an issue unless below a certain threshold

C. Do you see a functional difference between an evaluator answering “ultimate issue” questions regarding custody and stating who they perceive to be the better parent? Y/N

D. For each item listed below, please indicate how you feel on the 5-point Likert-type scale

	Poor Idea		Mixed Feelings		Very Good Idea
1. Joint physical custody as a choice	1	2	3	4	5
2. Joint legal custody as a choice	1	2	3	4	5

3. Sole physical custody as a choice	1	2	3	4	5
4. Sole legal custody as a choice	1	2	3	4	5

E. What is your preferred placement schedule: _____

F. What percentage of the time have you recommended physical custody of siblings be split between their parents: _____

G. At what age should children be able to choose to spend parenting time (“visit”) or not with a parent: _____

H. At what age should children be allowed to choose which parent they will live with: _____

I. Some states are enacting laws that suggest or require "substantially equal" parenting time for each parent. Please mark which of the following percentage splits you believe represent a "substantially equal" share of parenting time.

___50%/50%

___55%/45%

___60%/40%

___65%/35%

___70%/30%

___75%/25%

This is the end of the survey.

Thank you very much for your time and attention.

References

Abrams, T. A. (1988). The influence of family disruption on clinician bias in the assessment of children. Journal of Divorce, 11(3/4), 189 – 205.

Ackerman M. J. (2008). Does Wednesday mean mom's house or dad's? (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Ackerman, M. J. & Ackerman, M. C. (1997). Custody evaluation practices: A survey of experienced professionals (revisited). Professional Psychology: Research & Practice, 28, 137-145.

Ackerman, M. J., Ackerman, M. C., Steffen, L. J., & Kelley-Poulos, S. (2004). Psychologists' practices compared to the expectations of family law judges and attorneys in child custody cases. Journal of Child Custody, 1(1), 41-60

Ackerman, M. J. & Brey Pritzl, T. (2011). Child custody evaluation practices: A 20-year follow-up. Family Court Review, 49(3), 618-628.

Ackerman, M. J. & Steffen, L. J. (2001). Child custody evaluation practices: A survey of Family Law Judges. American Journal of Family Law, 14(1), 12-23.

American Professional Society on the Abuse of Children (1990). Guidelines for psychosocial evaluation of suspected sexual abuse in young children. Chicago, IL: Author.

American Psychological Association (1994). Guidelines for child custody evaluations in divorce proceedings. Washington, DC: Author.

Appel, A. E. & Holden, G. W. (1998). The co-occurrence of spouse and physical child abuse: A review and appraisal. Journal of Family Psychology, 12, 578-599.

Association of Family and Conciliation Courts (2003). AFCC Model Standards of Practice for Child Custody Evaluation. Madison, WI: Author.

Austin, W. G. (2002). Guidelines for utilizing collateral sources of information in child custody evaluations. Family Court Review 40(2), 177-184.

Austin, W. G., Jaffe, P. & Friedman, B. (1994). Custody and access assessors: Effects of background and experience on analogue case judgment. Canadian Journal of Behavioural Science, 26, 463-475.

Ayoub, C. C., Deutsch, R. M., & Maraganore, A. (1999). Emotional distress in children of high-conflict divorce: The impact of marital conflict and violence. Family and Conciliation Courts Review, 37, 297-314.

Bala, N. (1990). Assessing the assessor: Legal issues. Canadian Family Law Quarterly, 6, 179-226.

Beverly, D. P., & McSweeney, E. A. (1987). Social welfare and social justice. Englewood Cliffs, NJ: Prentice Hall.

Bexar County Domestic Relations Office (2013, Spring). Bexar County Civil District Court Social Study Evaluator List. (Available from the Bexar County Domestic Relations Office, 100 Dolorosa, 3rd Floor, San Antonio, Texas 78205).

Bow, J.N. (2006). Review of empirical research on child custody practices. Journal of Child Custody, 3(1), 23-50.

Bow, J. N. & Boxer, P. (2003). Assessing allegations of domestic violence in child custody evaluations. Journal of Interpersonal Violence, 18(12), 1394 – 1410.

Bow, J. N., Flens, J. R., Gould, J. W. and Greenhut, D. (2005). An analysis of administration, scoring, and interpretation of the MMPI-2 and MCMI-II/III in child custody evaluations. Journal of Child Custody, 2(4), 1-22.

Bow, J. N. & Quinnell, F. A. (2004) Critique of child custody evaluations by the legal profession. Family Court Review, 42(1), 115-127.

Bow, J. N. & Quinnell, F. A. (2002) A critical review of child custody evaluation reports. Family Court Review, 40(2), 164-176.

Bow, J. N. & Quinnell, F. A. (2001). Psychologists' current practices and procedures in child custody evaluation: Five years after American Psychological Association guidelines. Professional Psychology: Research & Practice, 32, 261-268.

Bow, J. N., Quinnell, F. A., Zaroff, M. & Assemany, A. (2002). Assessment of sexual abuse allegations in child custody cases. Professional Psychology: Research & Practice, 33, 566-575.

Bramlett, M. D., & Mosher, W. D. (2001). First marriage dissolution, divorce, and remarriage: United States [Advance data from Vital and Health Statistics Rep. No. 323]. Hyattsville, MD: National Center for Health Statistics.

Brandt, R., Dawes, A., Africa, A. & Swartz, L. (2004). A thematic content analysis of psychologists' reports in child custody evaluations. South African Journal of Psychology, 34(2), 259-282.

Brey, T. R. (2007). Child custody evaluation practices: A survey of psychologists. Unpublished doctoral dissertation, Wisconsin School of Professional Psychology.

Bureau of Labor Statistics. (2012). Labor Force Statistics from the Current Population Survey. Retrieved November 1, 2013 from http://www.bls.gov/cps/cpsaat11.htm

Caplan, P. J. & Wilson, J. (1990). Assessing the child custody assessors. In Steinberg, D. M. (Founding Ed.) Reports of Family Law (3rd series), Volume 27 (pp. 121 - 134). Toronto: Carswell.

Centers for Disease Control and Prevention. (2013). National Marriage and Divorce Rate Trends. Retrieved February 15, 2013 from http://www.cdc.gov/nchs/nvss/marriage_divorce_tables.htm

Clarke, S. C. (1995). Advance report of final divorce statistics, 1989 and 1990. (Monthly Vital Statistics Report, Vol. 43, No. 9, Supplement). Hyattsville, MD: National Center for Health Statistics.

Cohen, J. (1992). A power primer. Psychological bulletin Vol 112(1), 155-159.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, O. & Shnit, D. (2001). Social workers' recommendations on the non-custodial father's visitation rights with his preschool children. International Social Work, 44(3), 311-328.

Crosby-Currie, C. A. (1996). Children's involvement in contested custody cases: Practices and experiences of legal and mental health professionals. Law and Human Behavior, 20(3), 289-311.

Davidson-Arad, B., Cohen, O. & Wozner, Y. (2003). Social workers custody recommendations: Contributions of child's expected quality of life and parental features. Journal of Divorce & Remarriage, 39(1/2), 11-26.

DeClue, G. (2002). The best interests of the village children. Journal of Psychiatry & Law, 30, 335-390.

Denton County District Court Administrator. (2013). Approved List of Social Workers for Denton County. Retrieved March 23, 2013 from http://dentoncounty.com/dept/main.asp?Dept=32&Link=725

Dickerson, J. L., Allen, M & Pollack, D. (2011) How to screen adoptive and foster parents. Washington, DC: NASW Press.

Fisher, R.A. (2000). Mathematics of a Lady Tasting Tea. In J. Newman (Ed.), The world of mathematics, Volume 3 (pp. 1512 – 1521). Mineola, New York: Courier Dover Publications.

Fitzgerald, J. M. & Moltzen, N. (2004). Psychological evaluation of the child's best interests: the interpretation of data in the preparation of child welfare reports in the New Zealand family court. Psychiatry, Psychology and Law, 11(2), 214-226.

Flens, J. R. (2004, October). Advanced institute on psychological testing. Pre-symposium Institute conducted at the fourth annual Association of Family and Conciliation Courts Child Custody Symposium, Nashville, TN.

Gourley, E. V. & Stolberg, A. L. (2000). An empirical investigation of psychologists' custody evaluation procedures. Journal of Divorce & Remarriage, 33, 1-29.

Gravetter, F. J. & Wallnau, L. B. (2004). Statistics for the Behavioral Sciences. Belmont, CA: Thomson Learning.

Hagen, M.A. (1997). Whores of the court. New York: ReganBooks.

Hagen, M. A. & Castagna, N. (2001). The real numbers: Psychological testing in custody evaluations. Professional Psychology: Research and Practice, 32(3), 269-271.

Horvath, L. S., Logan, TK & Walker, R. (2002). Child custody cases: A content analysis of evaluations in practice. Professional Psychology: Research and Practice, 33(6), 557-565.

Hughes, D. S. & O'Neal, B. C. (1983). A survey of current forensic social work. Social Work, September-October, 393-394.

Jaffe, P. & Cameron, S. (1984). A preliminary study of clinical interventions in custody & access disputes. Canadian Journal of Behavioral Science, 16, 167-172.

Jameson, B. J., Ehrenberg, M. F., & Hunter, M. A. (1997). Psychologists' ratings of the best-iterests-of-the-child custody and access criterion: A family systems assessment model. Professional Psychology: Research and Practice 28(3), 253 – 262.

Keilin, W. G. & Bloom, L. J. (1986). Child custody evaluation practices: A survey of experienced professionals. Professional Psychology: Research & Practice, 17, 338-346.

Kunin, C. C., Ebbesen, E. B. & Konecni, V. J. (1992). An archival study of decision-making in child custody disputes. Journal of Clinical Psychology, 48(4), 564-573.

LaFortune, K. A. & Carpenter, B. N. (1998). Custody evaluations: A survey of mental health professionals. Behavioral Sciences & the Law, 16, 207-224.

Lamb, M. E. & Kelly, J. B. (2009) Improving the quality of parent-child contact in separating families with infants and young children: Empirical research foundations. In Galatzer-Levy, R.M., Kraus, L. & Galatzer-Levy, J. (Eds.) The Scientific Basis of Child Custody Decisions (pp. 187-214). Hoboken, NJ: John Wiley & Sons, Inc.

Lemon, N. K. D. (2000). Custody and visitation trends in the United States in domestic violence cases. Journal of Aggression, Maltreatment & Trauma, 3, 329-343.

Logan, TK, Walker, R., Jordan, C. E. & Horvath, L. (2002). Child custody evaluations and domestic violence: Case comparisons. Violence and Victims, 17(6), 719-742

Lowrey, C. R. (1985). Child custody evaluations: Criteria and clinical implications. Journal of Clinical Child Psychology 14(1), 35-41.

Lugaila. (1998). Marital status and living arrangements: May 1998 (update). U.S. Census Bureau Current Population Reports: Population Characteristics, Series P20-514.

Maccoby, E. E., & Mnookin, R. H. (1992). Dividing the child: Social and legal dilemmas of custody. Cambridge, MA: Harvard University Press.

Martindale, D. A., Tippins, T. M., Ben-Porath, Y. S., Wittmann, J. P. & Austin, W. G. (2012). Assessment instrument selection should be guided by validity analysis not professional plebiscite: Response to a flawed survey. Family Court Review, 50(3), 502-507.

McCannn, J. T., Flens, J. R., Campagna, V., Collman, P., Lazzaro, T., and Connor, E. (2001). The MCMI-III in child custody evaluations: A normative study. Journal of Forensic Psychology Practice, 1, 27-44

Milchman, M. S. (2000). Mental health experts' common error in assessing bonding in guardianship cases. Journal of Psychiatry & Law, 28, 351-378.

Monette, D. R., Sullivan, T. J. & DeJong, C. R. (2002). Applied social research: Tool for the human services (5^{th} ed.). Toronto, Ontario: Wadsworth Thomson Learning.

Murrie, D.C., Boccaccini, M.T., Guarnera, L.A. & Rufino, K.A. (2013). Are forensic experts biased by the side that retained them? Psychological Science, 24(10), 1889-1897.

Ostrom, B. J.,& Kauder, N. B. (1996). Examining the work of state courts, 1995: A national perspective from the court statistics project. Williamsburg, VA: National Center for State Courts.

Pruett, M. K. & Barker, C. (2009) Joint custody: A judicious choice for families – but how, when, and why? In Galatzer-Levy, R.M., Kraus, L. & Galatzer-Levy, J. (Eds.) The Scientific Basis of Child Custody Decisions (pp. 417-462). Hoboken, NJ: John Wiley & Sons, Inc.

Quinnell, F. A. & Bow, J. N. (2001). Psychological tests used in child custody evaluations. Behavioral Sciences and the Law, 19, 491-501.

Remley, T. P., Jr., & Miranti, J. G. (1991). Child custody evaluator: A new role for mental health counselors. Journal of Mental Health Counseling, 13, 334-342.

Robb, A. (2006). Strategies to address clinical bias in the child custody evaluation process. Journal of Child Custody 3(2), 45-69.

Rubin, A. & Babbie, E. (2008). Research Methods for Social Work (6th ed.). Belmont, CA: Thomson Brooks/Cole.

Sagi, A. & Dvir, R. (1993). Value biases of social workers in custody disputes. Children and Youth Services Review, 15, 27-42.

Schepard, A. I. (2004). Children, courts, and custody: Interdisciplinary models for divorcing families. Cambridge: Cambridge University Press.

Schindler, R. (1985). Court orders and custody evaluations in Israel. Child Welfare LXIV(4), 383-393.

Tippins, T. M. & Wittman, J. P. (2005). Empirical problems with custody recommendations: A call for humility and judicial vigilance. Family Court Review, 42(2), 193-222.

Uniform Child-Custody Jurisdiction and Enforcement Act (1997). 9(1A) U.L.A. 657 (1999). Available from the Uniform Law Commission (formerly National Conference of Commissioners on Uniform State Laws) at www.uniformlaws.org .

Ver Steegh, N. & Dalton, C. (2008). Report from the wingspread conference on domestic violence and family courts. Family Court Review, 46(3), 454-475.

Warshak, R. A. (1996). Gender bias in child custody decisions. Family & Conciliation Courts Review Vol 34, 396-409.

Warshak, R. A. (1992). The custody revolution: The father factor and the motherhood mystique. New York: Poseidon.

Zervopoulos, J. A. (2008). Confronting mental health evidence: A practical guide to reliability and experts in family law. Chicago: ABA Publishing

Biographical Information

Aaron D. Robb is a 1990 graduate of the Texas Academy of Mathematics and Science in Denton, Texas. He received his Bachelor of Science degree in Psychology in 1993 from Texas A&M University in College Station, Texas and his Master of Education degree in Counseling and Student Services in 1997 from the University of North Texas. He is a National Merit Scholar, a recipient of the Julian C. Stanly Student Life Scholarship, and recipient of the Texas A&M President's Endowed Scholarship. Dr. Robb was the 2006 Social Work Student of the Year for the Dallas Branch of the National Association of Social Workers, was inducted into the Phi Kappa Phi Honors Society in 2007, and was named a University Scholar at the University of Texas at Arlington in 2008. He has twice been named to Who's Who Among Students in American Universities and Colleges, first in 2007 and most recently in 2012.

Following completion of his first graduate degree Dr. Robb worked for Child Protective Services in the Dallas – Ft. Worth area while also serving as a volunteer clinician at the Denton County Children's Advocacy Center. It was in these settings where Dr. Robb experienced firsthand the complicated interplay of legal and social services and developed a life-long commitment to high quality services for children and families involved with the court system. In addition to his private practice work, he is the former President of the Texas State Chapter of the Association of Family and Conciliation Courts, an editorial board member of the Journal of Child Custody, and has been a regular presenter at national and international conferences. His ongoing research interests include child custody evaluation, court-connected mental health services, co-parenting programs, and the needs of children involved in the legal system.