

Running head: CONTENT AND QUALITY OF CHILD WELFARE EVALUATIONS

AN EXAMINATION OF THE CONTENT AND QUALITY OF FORENSIC EVALUATIONS
IN CHILD WELFARE

A DISSERTATION
SUBMITTED TO THE FACULTY
OF
THE GRADUATE SCHOOL OF APPLIED AND PROFESSIONAL PSYCHOLOGY
OF
RUTGERS,
THE STATE UNIVERSITY OF NEW JERSEY

BY
CHRISTINA N. DIAZ
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
OF
DOCTOR OF PSYCHOLOGY

NEW BRUNSWICK, NEW JERSEY

AUGUST 2020

APPROVED:

N. Andrew Peterson, Ph.D.

Elisa Shernoff, Ph.D.

DEAN:

Francine Conway, Ph.D.

CONTENT AND QUALITY OF FORENSIC EVALUATIONS

Copyright 2020 by Christina Diaz

Abstract

The field of forensic psychology is ever-growing, though there is still a lack of research that examines the content and quality of forensic assessments. Literature is even more limited when examining the subfield of child welfare. Forensic evaluations in child welfare are important as they inform treatment recommendations for members of vulnerable populations. The purpose of this study was to examine the relationship between the content (i.e., background information and psychological tests) and quality of forensic evaluations in child welfare, with the goal of contributing to literature to better inform practice in this field. This study used a sample of over 1600 evaluations reviewed using the Quality Improvement Tool (QI Tool). These data were originally collected by the New Jersey Coordination Center for Child Neglect Forensic Evaluation and Treatment (NJCC). The QI Tool was designed to adhere to the principles outlined in the *New Jersey Department of Children and Families (DCF) Guidelines for Expert Evaluations in Child Abuse/Neglect Proceedings (2012)*. The results of this study found that there is a statistically significant positive correlation between the variables of interest (Total Background Information and Total Tests) and the overall quality of the evaluation. This positive correlation remained when the sample was stratified by the age of the subject. Regression analyses were conducted to determine if the overall quality of the evaluation could be predicted by the number of unique background sources used in an evaluation and the number of psychological tests used in an evaluation. The results found that using a higher number of unique sources of background information and a higher number of psychological tests predicted higher levels of quality. Additionally, evaluations conducted with a child as the subject tended to use more unique sources of background information and psychological tests. Rank tests were also conducted to determine if there were group differences in the number of unique background

sources and number of psychological tests used based on whether the evaluator used a multimodal approach or demonstrated expertise in testing, respectively. The results found that there are statistically significant differences in the groups based on the aforementioned measures of quality. Limitations of this study and suggestions for future research are also discussed.

Acknowledgements

This dissertation would not have been possible without NJCC and its team of amazing researchers, especially Erica Lee and Matthew Forsythe. Thank you for both for your endless support over the past several years. Being part of NJCC truly changed my life for the better and I would not be where I am today without the valuable research experience I attained working as part of this team. I entered my role as a research assistant with an interest in child welfare and left with a passion for child welfare.

I want to thank my family for always stressing the importance of education. I am blessed to have been raised by so many intelligent, hardworking individuals who always encouraged me to be the best person I can be. Mom and dad, I love you both so much and hope that I've made you proud.

Thank you to Dr. N. Andrew Peterson and Dr. Elisa Shernoff for being an awesome and supportive committee. Your encouragement, mentorship, and guidance have been invaluable to me throughout this process. This dissertation has been one of the most challenging yet rewarding projects I've ever completed, and I could not have done it without the support of two excellent research mentors.

I would like to thank my friends as well. This has been an arduous, nearly two-year process and you all have stood by my side and offered support throughout, especially in the past few months as I was troubleshooting my analyses and writing up the results. I'm so looking forward to us getting to have fun again now that this process is almost over!

Finally, thank you to my cats, Duchess and Mortimer. Where would I be without your unconditional love and cuddles? Finding you two transformed me from a person who was apathetic about animals at best, to a certified cat lady and I would not have it any other way.

TABLE OF CONTENTS

ABSTRACT.....ii

ACKNOWLEDGEMENTSiv

LIST OF TABLESviii

INTRODUCTION1

LITERATURE REVIEW4

 Brief History of Forensic Assessment4

 Studies of the Content of Forensic Assessments5

 Child Custody Evaluations6

 General Forensic Assessment Procedures12

 The Role of Psychological Testing in Forensic Assessment13

 Studies of Psychological Testing Procedures in Forensic Assessment15

 The Role of Background Information in Forensic Assessment17

 Multimodal approaches to forensic assessment.....18

 Studies of the use of Background Information in Forensic Assessment18

 Quality Improvement in Forensic Assessment19

 Studies of the Quality of Forensic Assessments20

 The Current Study.....22

METHODS23

 Design23

 Procedures23

 Measures26

 Collateral Information.....26

Variable: Total Background Information.....	26
Psychological Inventories and Interpretation.....	27
Variable: Total Tests.....	28
Evaluation Rubric	28
Variable: Use of a multimodal approach	28
Variable: Expertise in testing.....	29
Variable: Overall Quality.....	29
Sample.....	29
Analytic Approach	29
RESULTS	30
Data Preparation.....	30
Analyses for Questions 1 and 2	31
Question 1.....	32
Question 2	34
Analyses for Questions 3 and 4	35
Question 3	35
Question 4.....	37
DISCUSSION.....	38
Background Information and Quality	40
Psychological Testing and Quality	42
Limitations of the Study.....	44
Implications for Research and Practice.....	47

Conclusion	48
REFERENCES	49
APPENDICES	57

LIST OF TABLES

Table	Page #
1. Descriptive Statistics for Questions 1 and 2	31
2. Descriptive Statistics for Overall Quality for Question 1	32
3. Summary of Logistic Regression Analysis for TBI and Age Predicting Overall Quality.....	34
4. Descriptive Statistics for Multimodal Approach	34
5. Descriptive Statistics for Questions 3 and 4	35
6. Descriptive Statistics for Overall Quality for Question 3.....	35
7. Summary of Logistic Regression Analysis for TT and Age Predicting Overall Quality.....	37
8. Descriptive Statistics for Level of Expertise	38

Introduction

An Examination of the Content and Quality of Forensic Evaluations in Child Welfare

In 2017, the New Jersey Division of Child Protection and Permanency (DCP&P) received a total of 87,574 reports for investigations of abuse and/or neglect (New Jersey Child Welfare Data Hub, 2018). Of these initial child welfare reports, 9,500 were investigated and one or more individuals involved in the case underwent forensic psychological or parenting capacity evaluations, broadly referred to as forensic evaluations in child welfare (Forsythe, 2018). For the purposes of this study, the term “forensic evaluation in child welfare” will be defined as it is in the *New Jersey Department of Children and Families Guidelines for Expert Evaluations in Child Abuse/Neglect Proceedings (Mental Health)* (2012). These guidelines indicate that a forensic evaluation in child welfare is considered the following:

A forensic evaluation in child welfare proceedings and child protective service matters is an evaluation necessary to assist the court and/or CP&P in case planning, or to resolve a case. A forensic evaluation may be requested by CP&P, by another party to a proceeding, or the court. Any evaluation that may reasonably be expected to be submitted to the court is termed forensic (New Jersey Department of Children and Families, p. 4).

These guidelines were created by a multidisciplinary panel of experts in child welfare. They address best practices in conducting forensic evaluations in child welfare and are intended to improve both the quality and utility of these evaluations (New Jersey Department of Children and Families, 2012).

The primary role of a forensic evaluation in child welfare is to assess a child’s degree of safety and functioning following an incident in which they were harmed or at risk of being harmed (New Jersey Department of Children and Families, 2012). Additionally, many forensic evaluations in child welfare include treatment recommendations based on the psychological,

developmental, and behavioral needs of the subject of the evaluation (New Jersey Department of Children and Families, 2012). In the state of New Jersey, the majority of forensic evaluations in child welfare are conducted or supervised by a licensed psychologist. However, there are limited circumstances where an independent evaluation by a licensed clinical social worker may be appropriate (New Jersey Department of Children and Families, 2018). Because these reports are typically consumed by professionals outside of the field of psychology (e.g., attorneys, judges, and case workers), it is imperative that high quality reports are produced where the findings are communicated in a way that is accessible to professionals outside of the mental health field (Otto, DeMier, & Boccaccini, 2014).

Hayes, Nelson, and Jarrett (1987) highlight the importance of clinical assessment on treatment outcomes. The authors used the phrase “the treatment utility of assessment” to describe the extent to which assessment contributes to beneficial treatment outcomes (Hayes et al., 1987, p. 961). Historically, assessment was seen as a “negative intrusion into the therapeutic alliance” and there was limited literature on the integration of assessment and treatment (Nelson-Gray, 2003, p. 521). Nelson-Gray (2003) suggested that further studies are needed to determine the extent to which assessment contributes to treatment outcomes.

Meyer et al. (2001) describe several purposes of psychological assessment, including as a means of identifying therapeutic needs and possible outcomes. Hodges (2004) has a similar point of view and describes the primary purpose of assessments as a means of facilitating the work of the treatment provider for the benefit of the recipients of services. Still, until recently, there were no articles that specifically examined treatment outcomes of forensic assessment. This changed in 2017, when a special issue of *Psychological Assessment* was published that examined the field utility of forensic assessment instruments and procedures. In the introduction of this issue, Edens

and Boccaccini (2017) focus on why field studies are important to improving the quality of forensic assessment. The authors argue that the findings of studies that solely rely on “lab” data (e.g. assessment data does not have “real-world implications” for the examinee) may not generalize to clinical practice due to differing administration procedures (Edens & Boccaccini, 2017, p. 601). Overall, this field would benefit from further research on specific treatment outcomes.

To the author’s knowledge, there is one empirical study that briefly discusses the content of forensic evaluations in child welfare including the use of background information (i.e., record reviews and collateral or, secondary, interviews with other parties) and psychological testing procedures (Neil & Grisso, 2014). There do not appear to be any studies that discuss the quality of forensic evaluations in child welfare. There are some studies that examine the content and quality of other types of forensic assessments, although not all findings are able to be generalized to the field of child welfare due to the differences in psycholegal questions that must be answered in various forms of forensic assessment (Ackerman & Ackerman, 1996; Ackerman & Brey Pritzl, 2011; Bow & Quinnell, 2002; Hecker & Steinberg, 2002; Heilbrun, Rosenfeld, Warren, and Collins, 1994; Keilin & Bloom, 1986; LaFortune & Carpenter, 1998; Lander & Heilbrun, 2009; Neal & Grisso, 2014; Petrella & Poythress, 1983; Quinnell & Bow, 2001).

The current study seeks to fill a gap in the literature about forensic evaluations in child welfare by examining the relationship between certain aspects of the content of forensic evaluations in child welfare (i.e., record reviews, collateral interviews, and psychological testing) and the overall quality of these evaluations. Currently, there is no consensus on the evaluative procedures that should be used in forensic evaluations in child welfare. As such, the findings of

this study may be used to make recommendations for the procedures of forensic evaluations in child welfare with the hopes of improving the quality and utility of these evaluations.

Literature Review

Brief History of Forensic Assessment

For the purposes of this study, the term “forensic assessment” will apply broadly to any psychological assessment that can be used to inform a legal decision. Until the 1960s, techniques used for forensic assessments typically did not differ from those used for traditional psychological assessments (Heilbrun, Rogers, & Otto, 2002). Specifically, psychologists did not structure their evaluative procedures around the referral questions which the courts sought to have answered (Heilbrun et al., 2002). Psychologists generally relied on clinical interviews and measures of intelligence, achievement, or personality as their primary sources of information; little emphasis was placed on the legal aspect of the assessment (Heilbrun et al., 2002). The first test created specifically for use in forensic assessments was published in 1965; A Checklist of Competency to Stand Trial was used to determine the degree to which an individual’s mental illness or intellectual capabilities affected their ability to understand the legal process (Heilbrun et al., 2002; Robey, 1965). Similar instruments were developed in the following decades.

The child welfare subfield of forensic psychology began to receive more attention in the 1980s. The Bricklin Perceptual Scales (BPS) were published in 1984 and serve as an early example of a psychological test specifically designed for child custody issues; this test is designed to assess a child’s perception of the quality of their parental relationship(s) (Bricklin, 1984). Although the BPS has various psychometric limitations, this measure remains an important part of the history of forensic assessment (Heilbrun, Rogers, & Otto, 2002). Another early example of a child welfare related psychological test is the Child Abuse Potential Inventory

(CAPI) (Milner, 1986). This measure seeks to assist child protection workers in investigating allegations of physical abuse (Laulik, Allam, & Browne, 2013; Milner, 1986). In the decades since the BPS and CAPI were published, dozens of measures for use in child welfare matters have been published.

In 1991, the American Psychological Association created the Specialty Guidelines for Forensic Psychology (APA, 2013b). Eight years later, the APA created the Guidelines for Psychological Evaluations in Child Protection Matters (2013a). The APA has also created the Guidelines for Child Custody Evaluations in Family Law Proceedings (2010). Each set of guidelines are informed by the APA's Ethical Principles of Psychologists and Code of Conduct (2010; 2013a; 2013b). The Specialty Guidelines for Forensic Psychology and Guidelines for Psychological Evaluations in Child Protection Matters were both updated in 2013 to reflect continuing developments in professional practice. It is important to note that the term "guidelines" only refers to suggestions or recommendations for professional behavior; neither of these sets of guidelines is mandatory (APA, 2010, 2013a, 2013b). Currently, there are no established standards of practice or standards of care for forensic evaluations in child welfare.

Studies of the Content of Forensic Assessments

There appears to be one extant study that examines the content of forensic evaluations in child welfare (Neal & Grisso, 2014). However, there are various studies that examine the content of other types of forensic assessments (Ackerman & Ackerman, 1996; Ackerman & Brey Pritzl, 2011; Archer, Buffington-Vollum, Vauter Stredny, & Handel, 2006; Bow & Quinnell, 2002; Keilin & Bloom, 1986; LaFortune & Carpenter, 1998; Lees-Haley, 1992; Neal & Grisso, 2014; Quinnell & Bow, 2001). One early example of a study examining the content of forensic assessments is Keilin and Bloom's (1986) survey of child custody evaluation practices. It is

important to note, however, that the child custody evaluations discussed in Keilin and Bloom's (1986) article and subsequent similar articles are somewhat different from the previously defined forensic evaluations in child welfare. The most notable difference is that child custody evaluations are typically conducted to assist in custody and visitation planning between parents and/or legal guardians who are in the process of divorcing (Stahl, 1999). Unlike forensic evaluations in child welfare, these child custody evaluations are not necessarily the result of a child experiencing harm or a child being at risk of harm.

Child Custody Evaluations. Keilin and Bloom's (1986) survey included 82 mental health practitioners selected from directories of forensic experts. In this study, 100% of providers stated that they conducted clinical interviews with parents, 98.8% conducted clinical interviews with children, 75.6% of respondents endorsed psychological testing of parents, and 74.4% endorsed psychological testing of children. Just under one half (48.8%) of providers engaged in contact with collateral sources, spending approximately 1.32 hours on this activity. There is no information on the percentage of providers who review records nor the amount of time spent on this activity. The authors stated that 12.2% of providers endorsed engaging in "other" activities for approximately 1.96 hours, but there is no specific information on what "other" could entail and whether that captures the missing information about reviewing records (Keilin & Bloom, 1986, p. 340). In total, providers reported spending 18.8 hours on all procedures of child custody evaluations. The majority of respondents who endorsed using psychological tests reported using the Minnesota Multiphasic Personality Inventory (MMPI) for adult clients. Other tests used with adult clients (in descending order of percentage of respondents using the test) included the Rorschach (41.5%), the Thematic Apperception Test (TAT) (37.8%) the Wechsler Adult Intelligence Scale (WAIS) (29.3%), and the Bender Visual Motor Gestalt Test (Bender-Gestalt)

(12.2%) (Keilin & Bloom, 1986). There was no single test that the majority of respondents reported using on children, however, 45.1% of respondents reported using some form of cognitive assessment with children (Keilin & Bloom, 1986). In descending order of reported use, respondents also used the TAT or Children's Apperception Test (CAT) (39.0%), unspecified projective drawings (32.9%), the Rorschach (29.2%), and the Bender-Gestalt (23.2%). In the conclusion, Keilin and Bloom (1986) state that though adults are typically administered the MMPI, they may also be administered projective and cognitive assessments when appropriate. However, there is no further information on what referral questions might deem the use of these tests appropriate. This article served as a model for examining the content of forensic assessments, though the lack of information on record reviews makes it difficult to determine the degree of importance that this activity has in the assessment process.

Ackerman and Ackerman (1996) replicated Keilin and Bloom's (1986) study and expanded it by incorporating new questions about child custody evaluation practices based on professional experiences and literature reviews of the topic. This study found that the average amount of time spent on evaluation procedures increased from 18.8 hours to 26.4 hours, with providers spending more time writing the report and reviewing records. The Ackerman and Ackerman (1996) study shows that providers reported spending 2.6 hours reviewing records and 1.6 hours interviewing collateral sources. The authors of this study did not report the percentages of forensic evaluators that endorsed engaging in these activities. In regard to psychological testing, 98% of evaluators used psychological tests with adults and 92% used psychological tests with children, a fairly substantial increase from the 75.6% and 74.4% reported by Keilin and Bloom (1986), respectively (Ackerman & Ackerman, 1996). On average, 4.8 psychological tests were administered to children and 4.5 were administered to adults; similar figures were not

reported in Keilin and Bloom's (1986) study. The top two most popular tests for each group remained the same as in Keilin and Bloom's (1986) study, though Ackerman and Ackerman (1996) found that the most popular tests were used more often than in the previous study. The researchers found that 58.2% of evaluators used some form of cognitive assessment with children. Additionally, the MMPI was used by 91.5% of respondents for adult clients (Ackerman & Ackerman, 1986). Notably, there was a large increase in the use of the Millon Clinical Multiaxial Inventory (MCMI) with adult clients, whereas it was not mentioned in Keilin and Bloom's (1986) study. However, Ackerman and Ackerman (1996) state that the MCMI should be used with caution in child custody evaluations because it is designed for use with clinical populations. The inclusion of information about record reviews is valuable in that both the authors and practitioners who responded to the survey acknowledge the importance of reviewing records for forensic assessments.

LaFortune and Carpenter (1998) further expanded on the work of Ackerman and Ackerman (1996) and Keilin and Bloom (1986), though this survey used a five point Likert-type scale to report on child custody evaluation practices rather than determining the percentage of how many providers use each source of data. Additionally, the contribution of each evaluation procedure to the overall findings of the evaluation was also reported. On average, respondents reported spending 21.1 hours on all evaluation procedures, though there was high variability in the sample (LaFortune & Carpenter, 1998). Evaluators reported spending approximately 2.5 hours interviewing parents and rated these interviews as the most important components of child custody evaluation (LaFortune & Carpenter, 1998). Forensic evaluators stated that they spent about an hour each on interviewing collateral sources and reviewing school records, though the overall contribution was rated as just moderately important (LaFortune & Carpenter, 1998). The

MMPI was found to almost always be used as part of child custody evaluations and had a moderate contribution to the outcome of the case (LaFortune & Carpenter, 1998). Meanwhile, the Rorschach was rarely used and was considered to have little influence on the outcome of the case (LaFortune & Carpenter, 1998). However, it is difficult to compare these results to previous studies as tests used with children are not differentiated from tests used with adults.

In 2001, Quinnell and Bow conducted a survey of child custody practices in which participants were asked to rank child custody procedures in order of importance in addition to reporting the psychological tests they use. Psychological testing of parents was the fourth most important procedure out of ten and approximately 90% of parents were administered psychological tests (Quinnell & Bow, 2001). Psychological testing of children was the sixth most important procedure approximately 60% of children were administered psychological tests. Review of documents was ranked seventh, collateral contacts with school or health care providers was ranked eighth, and collateral contacts with spouses or relatives was ranked ninth (Quinnell & Bow, 2001). The number of children undergoing psychological testing is lower than previous studies, which Quinnell and Bow (2001) mention in their discussion of the findings. The study showed that the use of intelligence measures in custody evaluations declined compared to previous studies. In Quinnell and Bow's (2001) study, approximately half of all providers endorsed using intelligence measures with children and adults, though they only reported using these tests in approximately 30% of cases. There was a sevenfold increase in the use of parent rating scales such as the Achenbach Child Behavior Checklist (CBCL) and the Conners Parent Rating Scale (CPRS). It is suggested that the increased use of parent rating scales may reduce the need to directly test children (Quinnell & Bow, 2001). For adults, the MMPI continued to be administered in the majority of cases, with 91% of respondents endorsing its use

(Quinnell & Bow, 2001). The use of parenting inventories increased, with over 40% of providers endorsing the use of the Parent-Child Relationship inventory (PCRI) and the Parenting Stress Index (PSI) compared to just 11% and 9%, respectively, in Ackerman and Ackerman's (1996) study. Overall, objective tests were used more frequently with adults while projective tests were more frequently used with children (Quinnell & Bow, 2001). The popularity of the MCMI also continued to increase compared to prior surveys; it was found to be the second most commonly used test in this survey. However, the authors mention Rogers, Salekin, and Sewell's (1999) claim that the third edition of the MCMI may not meet the *Daubert* standard of admissibility in court due to low criterion and construct validity. The *Daubert* standard assesses whether expert testimony is scientifically valid and is relevant to the legal question (*Daubert v. Merrell Dow Pharmaceuticals, Inc., 1993*). Overall, this study highlights the importance of psychological testing for adults in forensic assessment, although records reviews and collateral interviews do not have as much emphasis placed on them.

Bow and Quinnell (2002) conducted a critical review of child custody evaluation reports addressing some of the limitations of survey research (e.g., using retrospective estimates regarding practices which may not accurately reflect the frequency of custody practices). In this study 100% of evaluations included interviews with the parents and 90.4% of evaluations included testing of both parents (Bow & Quinnell, 2002). Document reviews were included in 78.8% of evaluations. Collateral contacts with relatives, school personnel, doctors, therapists, and/or other significant figures were included in over 50% of reports. Finally, just 38.5% of evaluations included psychological testing of children (Bow & Quinnell, 2002). The majority of evaluators included in this study (87.8%) used objective personality tests with adults (Bow & Quinnell, 2002). Consistent with the aforementioned studies, the MMPI remained the most

popular test used, representing 93% of the sample; the MCMI was used in 44% of cases sampled. Parenting inventories were used in 44.9% of cases and the preferred instruments in this category were the PSI and PCRI (Bow & Quinnell, 2002). Within this study, projective personality tests (21.6%) were slightly more popular than objective personality tests (19.6%) for children. Preferred projective measures included the Rorschach, the Roberts Apperception Test, and unspecified family-themed projective drawings. Preferred objective measures included the Millon Adolescent Clinical Inventory (MACI) and the MMPI-Adolescent Version (MMPI-A) (Bow & Quinnell, 2002). This study is particularly relevant to the current study as it uses a similar method of data collection to determine the procedures used in the evaluation. The current study's method of data collection will be discussed in more detail later in this paper.

A follow up of Ackerman and Ackerman (1996) was conducted by Ackerman and Brey Pritzl (2011) in which major shifts in psychological test use were identified. For example, psychologists in Ackerman and Brey Pritzl's (2011) study spent approximately 46.1 hours engaging in all activities of a child custody evaluation, a vast increase from figures reported in previous studies (Ackerman & Ackerman, 1996; Keilin & Bloom, 1986). When compared to Ackerman & Ackerman (1996), the amount of time spent reviewing records increased from 2.6 hours to 5.6 hours while the time spent engaging in collateral contacts doubled from 1.6 hours to 3.2 hours. The amount of time spent on psychological testing also increased from 5.2 hours to 6.1 hours. Similarly to the previously mentioned studies, the MMPI remained the most commonly used psychological test for adults with 97.2% of psychologists reporting using the test (Ackerman & Brey Pritzl, 2011). The MCMI continued to rise in popularity with 71.3% of psychologists using the test. The MMPI-A was reportedly used by 66.2% of psychologists who evaluated children (Ackerman & Brey Pritzl, 2011). Notably, there was a substantial increase in

use of the CBCL, with 58.1% of psychologists in the sample using it versus just 4% in the Ackerman and Ackerman (1996) study. Projective measures continued to be popular among children and adult examinees, with 51.9% of psychologists reporting that they use the Rorschach on adults and 57.1% reporting that they use projective drawings with children (Ackerman & Brey Pritzl, 2011). This article serves as an excellent example of how procedures in forensic assessments evolved in terms of the amount of time spent completing all assessment procedures.

General Forensic Assessment Procedures. Neal and Grisso (2014) conducted an international survey of procedures in various forms of forensic assessment, including forensic evaluations for child welfare. The authors found that 100% of “child protection” evaluations (as they are referred to in this study) included an “examinee interview” (Neal & Grisso, 2014, p. 1411). The majority of child protection evaluations also used mental health and/or medical records, observations of the examinee, judicial records, and non-professional collateral interviews (e.g., grandparent or neighbor) to gather more information about the case (Neal & Grisso, 2014). Similarly, 100% of child custody evaluations included an examinee interview. The majority of child custody evaluations also used mental health/medical records, non-professional collateral interviews, and professional collateral interviews (e.g., therapist or lawyer) (Neal & Grisso, 2014). This study also examined the frequency of use of psychological tools in child protection evaluations. Consistent with previous findings, the vast majority of child protection (92.6%) and child custody evaluations (79.1%) used at least one structured psychological tool; on average, child custody evaluations used 3.77 tools and child protection evaluations used 4.65 tools, with a range of one to nine psychological tests used for both forms of forensic assessment (Neal & Grisso, 2014). The MMPI and MCMI were the most commonly used tools for both child custody and child protection matters; the Personality Assessment Inventory (PAI), PSI, CAPI,

and Rorschach complete the six most frequently used psychological tests for both types of evaluation (Neal & Grisso, 2014). This study serves as another example of the importance of using multiple sources of data to draw conclusions.

The Role of Psychological Testing in Forensic Assessment

Psychological testing is an area of expertise that differentiates psychologists from psychiatrists, social workers, professional counselors, and other mental health professionals. As a result, psychological testing typically plays a large role in a psychologist's daily practice (Otto, Edens, & Barcus, 2000). Psychological tests serve as an indicator of examinee behavior in regard to a specific construct; after testing data are obtained, they are evaluated and scored on a standardized scale (American Educational Research Association, APA, & National Council on Measurement in Education, 2014).

Within the context of forensic assessment, psychological testing is often seen as an objective and unbiased measure of assessment rather than solely relying on "clinical judgment" during interviews (Brodzinsky, 1993, p. 216). Past studies have shown that structured tools perform better than "unaided clinical judgment" in decision-making tasks (Neal & Grisso, 2014, p. 1407). Neal and Grisso (2014) identified three common reasons for using structured tools during forensic assessments. Providers reported that they used structured tools to have evidence-based methods of assessment, to improve credibility of the assessment, and to "standardize the assessment" (Neal & Grisso, 2014, p. 1414). Various structured assessment tools have been developed to assist in answering forensic referral questions; however, the use of these tools has been controversial, as some professionals have argued that these methods are too rigid and do not take into account constructs that are not measured by the tests (Neal & Grisso, 2014).

The role of psychological testing in forensic assessment is still unclear (Heilbrun, 1992; Neal & Grisso, 2014); however, surveys of psychologists have shown that the majority of those who conduct forensic assessments use at least one psychological tool as part of the assessment procedures (Ackerman & Ackerman, 1996; Ackerman & Brey Pritzl, 2011; Hagen & Castagna, 2001; Keilin & Bloom, 1986; Lally, 2003; Neal & Grisso, 2014; Quinnell & Bow, 2001). Lally (2003) emphasizes the importance of psychological tests as a third source of information alongside clinical interviews and collateral sources, though it is suggested that not all psychological tests are appropriate for the legal questions that may be raised by forensic assessments. The *New Jersey Department of Children and Families Guidelines for Expert Evaluations in Child Abuse/Neglect Proceedings (Mental Health)* (2012) state that evaluations should include a comprehensive record review, clinical interview, and appropriate psychological tests. Still, there are no specific recommendations for the use of psychological tests in forensic evaluations in child welfare.

To demonstrate expertise in psychological testing, the evaluator should select tests that are relevant to the legal issue and reliable (Heilbrun, 1995). Many psychologists are likely to rely on a handful of tests on which they were trained in graduate school (Cates, 1999); however, it is important that the tests selected are tailored to the referral question at hand (Heilbrun, 1992). Any psychological tests used should be well-documented in terms of development, administration procedures, and psychometric properties. Psychologists should also strive to select tests with a reliability coefficient of at least .80. If the reliability coefficient is less than .80, the psychologist should provide an explanation of why they selected the test, its strengths, and its limitations (Heilbrun, 1995).

Additionally, other important factors in demonstrating expertise in psychological testing are administration and interpretation. Whenever possible, tests should utilize standardized administration procedures (i.e., using the test in the condition in which it was normed) (Heilbrun, 1995). In the event that standardized testing procedures are not used, the psychologist should discuss the limitations of the administration. The generalizability of the results should also be discussed as the subject of the evaluation may be of a different population than the one on which the test was normed (Heilbrun, 1995). Response style is also a key aspect of interpretation as unusual response styles may lower the validity of the results (Heilbrun, 1995).

Studies of Psychological Testing Procedures in Forensic Assessment

Surveys of psychological testing procedures are relatively common in the field of psychology, although there have been fewer efforts to examine testing procedures in forensic assessment versus other subfields of psychology. Even fewer efforts have been made to study testing procedures in forensic evaluations for child welfare; the only extant study that appears to examine testing procedures for these evaluations was discussed previously in this literature review and is part of a larger study on the content of various forms of forensic assessment (Neil & Grisso, 2014). An examination of psychological test use in forensic assessment can serve as an informational tool for psychologists and training programs about widely accepted measures in the field (Lally, 2003). Within the context of the legal system, surveys of practices in forensic psychology that include data on test use can be used to argue whether or not a test meets criteria for admissibility in court (Lally, 2003). As the literature on testing procedures in forensic evaluations for child welfare is limited, surveys examining testing procedures in other forensic assessments will be discussed in this section.

In 1985, Lubin, Larsen, Matarazzo, and Seever suggested surveying psychological test use in specific settings; in response to this suggestion, Lees-Haley (1992) examined testing procedures in forensic assessment by administering a survey to 69 attendees of the American College of Forensic Psychology Annual Symposium. Participants were given a questionnaire with a list of psychological tests and were asked to rate their use of the test on a five-point Likert-type scale (Lees-Haley, 1992). Similar to Keilin and Bloom's (1986) findings, the most commonly used tests in order of popularity were the MMPI, WAIS, Rorschach, Bender-Gestalt, and unspecified sentence completion tasks (Lees-Haley, 1992). However, the results are limited by the lack of inclusion of the specific forensic settings in which these tests are used as well as the lack of operational definitions of the terms to judge frequency of use (Lees-Haley, 1992).

In 2006, Archer et al. conducted another survey of psychological test use in forensic assessment. The participants were asked to report on their use of psychological tests on a seven point Likert-type scale; weighted scores were then calculated for the total number of times a test was mentioned by psychologists in the sample and the frequency the tests were used by each respondent (Archer et al, 2006). Consistent with previously conducted studies, the MMPI had the highest weighted scores for multiscale inventories used with adult clients; the MMPI was used at least 50% of the time by all providers who reported incorporating it into their evaluations (Archer et al., 2006). The PAI and MCMI had the second and third highest weighted scores, respectively, demonstrating an increase in the acceptance of more recently developed multiscale inventories (Archer et al., 2006). Other popular tests that were reported as being used often in prior studies include Wechsler-branded measures of intelligence, unspecified sentence completion tasks, and the Rorschach (Archer et al., 2006). Additionally, this survey is unique in that it included a category for psychological tests used in child-related forensic issues. However,

the reason for referral for this category is not specified further, so it is not clear whether the evaluators were studying forensic evaluations for child welfare or other types of forensic assessment (e.g., juvenile competency to stand trial) (Archer et al., 2006). Still, similarly to the other surveys, the three most commonly used tests in this category were the MMPI-A, the PSI, and the CBCL (Archer et al., 2006).

The Role of Background Information in Forensic Assessment

As the aforementioned literature on the content of forensic assessments shows, forensic evaluators spend a substantial amount of time conducting interviews with collateral sources and reviewing records. For the purposes of this study, “background information” refers to information gained from third-party sources. Consistent with previous sources, this includes documents, records, and collateral interviews (Heilbrun, NeMoyer, King, & Galloway, 2015). Background information carries more weight in forensic assessments as clinicians may rely less on an examinee’s self-report than in typical clinical assessments since the examinee’s interests are not necessarily considered (Heilbrun et al. 2015). Specifically, background information can fill in gaps of self-report information and increase the reliability of forensic psychological evaluations (Heilbrun et al., 2015).

It is best practice to document all sources of data considered in the formulation of an opinion; the reader of the report should have an understanding of all sources of information on which opinions are based (Otto et al., 2014). Additionally, sources that were requested but ultimately were not available for review should be listed in the report. Any background information included in the final report should be relevant to the referral question (Otto et al., 2014).

Multimodal approaches to forensic assessment. Forensic practitioners should avoid relying on one source of data whenever possible and in the event that data cannot be corroborated, they should explain the limitations of their evaluation (AERA, APA, & NCME, 2014). The APA's Guidelines for Psychological Evaluations in Child Protection Matters (2013a) states that using multiple sources of data has three primary purposes. These purposes include broadening the evaluator's information base for opinions and recommendations, challenging biases that may compromise an evaluator's opinions and recommendations, and contributing to creating a quality evaluation that supports ethical and legally reliable opinions (APA, 2013a). Examples of multiple methods include clinical interviews, collateral interviews, clinical observations, record reviews, and psychological testing (APA, 2013a). In his seminal book, *Principles of Forensic Mental Health Assessment*, Heilbrun (2001) also highlights an individual's response style as an area in which multiple sources of data are beneficial as self-reported symptoms and behaviors can be more accurately assessed when third-party information is available. The studies described in this literature review suggest that the importance of using multiple sources of data is recognized by forensic practitioners.

Studies of the Use of Background Information in Forensic Assessment

To the author's knowledge, there is one extant study that examines the use of background information in forensic assessments. Heilbrun et al. (1994) sampled 593 Competency to Stand Trial (CST) and legal sanity evaluations from Florida and Virginia. These evaluations were conducted in hospital and community settings. The results of this study found that documentation regarding the subject of the evaluation's offense was included in the majority of hospital-based evaluations in both states and the majority of community-based evaluations in Virginia (Heilbrun et al., 1994). Mental health records were used much less frequently and the evaluators attribute

this to an “accessibility effect” (Heilbrun et al., 1994, p. 403). That is, evaluators may have had more difficulty accessing records that came from external sources (e.g., an evaluator in a hospital may have had more difficulty accessing mental health records from a different treatment provider) (Heilbrun et al., 2014). Although a practitioner’s ability to access records may vary depending on the case, it is still important to ensure that any attempts that were made to obtain documentation or conduct collateral interviews are documented.

Quality Improvement in Forensic Assessment

In its *Criteria for the Evaluation of Quality Improvement Programs and the Use of Quality Improvement Data*, the APA uses the Institute of Medicine’s definition of quality as “the degree to which services and treatment increase the likelihood of desired outcomes and are consistent with current professional knowledge” (APA, 2009, p. 551). Quality improvement in forensic assessment is an area that remains largely unexplored, although Heilbrun’s (2001) book is often looked to as a source for a model of forensic assessment. Heilbrun’s model has been adapted into a quality improvement instrument used in studies that examine several types of forensic assessments (Fuger, Acklin, Nguyen, Ignacio, & Gowensmith, 2014; Lander & Heilbrun, 2009; Nguyen, Acklin, Fuger, Gowensmith, & Ignacio, 2011; Sanschagrín, 2006). Because there is no set definition for quality regarding forensic evaluations in child welfare, the APA’s aforementioned definition will be used to define quality for the purposes of this study.

High quality evaluations are important because they reflect the examiner’s adherence to ethical standards and laws, focus on relevant legal questions, and communicate findings clearly and concisely (Nicholson & Norwood, 2000). High quality reports play an important role in establishing standards of practice for the field of child welfare. Improving the quality of these evaluations allows for the establishment of a standard of care (Heilbrun, Dematteo, Marczyk, &

Goldstein, 2008). Unlike standards of practice, professional standards of care are established by the court rather than members of the field itself. Violating professional standards of care may result in legal liability in addition to consequences levied by professional organizations (Heilbrun et al., 2008). For the purposes of this paper, it can be theorized that conducting and producing high quality forensic evaluations in child welfare will have a positive effect on outcomes for children and families with DCP&P involvement.

Studies of the Quality of Forensic Assessments

To the author's knowledge, there are no extant studies that quantitatively examine the quality of forensic evaluations in child welfare. However, there are a few studies that quantitatively examine the quality of other types of forensic assessments. One notable example of a study that fits these criteria was conducted by Petrella and Poythress (1983). The authors examined the quality of competency to stand trial and legal insanity evaluations conducted by psychologists, psychiatrists, and social workers at a public psychiatric hospital in Michigan (Petrella & Poythress, 1983). The authors examined evaluations for their level of detail which included the number of collateral interviews conducted, the number of requests for records from other sources, and the length of clinical notes (Petrella & Poythress, 1983). The authors then recruited an attorney, judge, and law professor to judge the quality of each report based on how clearly it presented information. These legal professionals judged the quality of these reports using a nine-question rating scale. The findings indicated the raters preferred reports that used multiple, external sources of information (Petrella & Poythress, 1983). As such, practitioners should strive to ensure that they are producing high quality reports by drawing conclusions from multiple sources of data.

Hecker and Steinberg (2002) quantitatively rated the content and quality of juvenile predisposition reports using a three-point Likert scale. Two raters (an undergraduate student majoring in psychology and a doctoral student in clinical psychology) were asked to note whether a specific content area was present in the report and whether or not the information provided was sufficient enough to answer the referral question and justify any recommendations made. The sufficiency of the eight content areas examined ranged from 10% to 63% (Hecker & Steinberg, 2002). The findings of this study indicated that many of the content areas examined in this study did not have sufficient information to justify the recommendations made by the evaluator, thus highlighting the necessity of practitioners meticulously gathering data to draw conclusions about a case.

A dissertation completed by Sanschagrin (2006) was the first to adapt Heilbrun's (2001) principles into a measure that assesses quality. The results of this dissertation found that many evaluators did not adhere to the principles of forensic mental health assessment, especially those regarding the formation of opinions, resulting in lower quality evaluations. Nguyen et al. (2011) and Fuger et al. (2014) used the same measure on adult criminal forensic evaluations, finding that reports still lacked quality with poor inter-rater reliability between psychologists who were rating the evaluations. The results of Nguyen et al. (2011) and Fuger et al. (2014) are important as they highlight concerns about the inter-rater reliability during quality improvement evaluations, a topic that was not discussed in earlier papers.

Lander and Heilbrun (2009) also examined the quality of forensic assessments in relation to their adherence to Heilbrun's (2001) outlined principles. A "Blue Ribbon Panel" (Lander & Heilbrun, 2009, p. 117) consisting of a judge, a law professor, an attorney, a psychiatrist, and a psychologist rated the quality of criminal forensic evaluations using a measure based on the one

used in Petrella and Poythress's (1983) study. The findings of this study indicated that longer reports were also deemed to be of higher quality. Additionally, forensic assessment reports that adhered to more of the outlined principles were rated more highly in terms of quality (Lander & Heilbrun, 2009). The reader is referred to Heilbrun's *Principles of Forensic Mental Health Assessment* (2001) for more specific information on the principles discussed in Lander and Heilbrun's paper (2009).

A retrospective study of juvenile forensic assessment reports was conducted in the Netherlands using a specially developed quality evaluation instrument called the Standardized Assessment Instrument of Reports (STAR) (Duits, van der Hoorn, Wiznitzer, Wettstein, & de Beurs, 2012). The STAR was developed to describe the usability of juvenile forensic evaluations; it consists of seven domains and allows for an objective score of quality or utility of the report (Duits et al., 2012). The results of this study found that the overall quality of evaluation reports increased significantly between 2005 and 2007, though reports were still lacking in areas measuring consistency, bias, and whether the referral questions were answered (Duits et al., 2012). This study is the first of its kind in that it used an objective measure to assess the overall quality of evaluation reports, similar to the measure used in the current study.

The Current Study

The current study sought to fill a gap in the aforementioned literature by examining the effect that the content of a forensic evaluation in child welfare has on its overall quality. This study uses extant data from a grant funded by the New Jersey Department of Children and Families; the overall goal of the larger study is to make recommendations for best practices in forensic evaluations in child welfare by reviewing selected evaluations for quality improvement purposes. Each evaluation was quantitatively measured for its quality by a reviewer with

expertise in forensic evaluations for child welfare. The current study examined the following research questions:

1. Is there a relationship between the number of unique sources of background information (i.e. record reviews and collateral information) used in a forensic evaluation in child welfare and the overall quality of the evaluation?
2. Does the number of unique sources of background information used in a forensic evaluation in child welfare differ based on the degree to which a multimodal approach was used to draw conclusions about the current case?
3. Is there a relationship between the number of psychological tests used in a forensic evaluation in child welfare and the overall quality of the evaluation?
4. Does the number of psychological tests used in a forensic evaluation in child welfare differ based on the level of expertise displayed by the evaluator?

Methods

Design

This study was correlational in nature as it examined the relationship between specific aspects of forensic evaluations in child welfare and the overall quality of evaluations. The data required for both research questions was previously collected, thus, this study qualifies as secondary data analysis. The approach was data-driven, as the research questions were designed based on the available data (Cheng & Phillips, 2014).

Procedures

The data used for the current study is part of a larger study funded by the New Jersey Department of Children and Families. The New Jersey Coordination Center for Child Abuse and Neglect Forensic Evaluation and Treatment (NJCC) seeks to inform best practices in child

welfare around the state, thereby improving service delivery and outcomes for children and families (Rutgers School of Social Work Institute for Families, n.d.). Thus far, NJCC has sampled approximately 16% of all forensic evaluations in child welfare conducted in New Jersey from 2015 to 2017. Forensic evaluations in child welfare that are included in this sample were conducted by private practitioners, non-profit agencies, and state funded Regional Diagnostic and Treatment Centers. NJCC has categorized evaluations into five catchment areas. Demographic data from the completed catchment areas show that the majority of forensic evaluations in child welfare were conducted on adult females.

This study was exempt from full review from the Institutional Review Board at Rutgers University under Exempt Categories four and five. This study used the NJCC Quality Improvement Tool (QI Tool) to collect information about the content and quality of forensic evaluations in child welfare. The QI Tool was designed by a team of multidisciplinary professionals employed by NJCC. It was then refined by a panel of psychologists with expertise in child welfare. The QI Tool is housed by Qualtrics and it includes 13 permanent blocks of questions that ask about demographic information of the subject of the evaluation, assessment procedures, recommendations, and the overall quality of the evaluation, among other topics. The QI Tool is currently on its third version and has been electronically distributed to a team of experts, referred to as peer reviewers, who have been contracted by NJCC as part of the general quality improvement study.

The peer review process seeks to enhance standards of practice by continuously providing feedback on services (Aimola et al., 2016; Grol, 1994). The majority of peer reviewers in this study are licensed psychologists with experience conducting forensic evaluations in child welfare for New Jersey DCF. Graduate student assistants on the NJCC research team are also

considered peer reviewers after their reviews have been deemed reliable by the Research Project Coordinator. Of note, the author of the current study is considered to be a peer reviewer for NJCC.

If accepted to the peer reviewer program, psychologists undergo one day of training in which they review the expectations for this project and complete quality reviews for sample cases. Peer reviewers (n=12) have monthly phone conferences with NJCC staff to ask questions or express concerns they might have about the project. The peer reviewer program is influenced by the theories guiding Participatory Action Research; quality improvement relies on the knowledge of peer reviewers to affect change in forensic psychological practice (Brydon-Miller, 1997). After training is complete, peer reviewers are also given an electronic version of the NJCC Codebook which contains information about how to assess each aspect of forensic evaluations in child welfare. The Codebook was developed to ensure consistency among the reviewers. It outlines each section of the QI Tool and provides examples that peer reviewers can use to guide their clinical judgment in regard to assigning ratings of quality. Ratings of quality for specific aspects of each evaluation and an overall rating of quality are generated independently.

Peer reviewers have the option to request a secondary review if they are unsure about the quality of an evaluation and would like a second opinion. A peer reviewer's first ten evaluations are automatically reviewed for reliability. Additionally, approximately 11% of the evaluations were randomly selected from each catchment area for secondary review to assess the reliability of the measure. The intraclass correlation coefficient (ICC) was used to assess inter-rater reliability. The ICC can range from zero to one with scores closer to one indicating higher similarity among groups (Koo & Ti, 2016). The third version of the QI Tool was determined to

have good reliability (ICC = .824). This is a substantial increase from the second version of the QI Tool which had moderate reliability (ICC = .680). The reliability for the first version of the QI Tool is not available.

Measures

The QI Tool contains 13 permanent sections of questions. Occasionally, a fourteenth section is added for exploratory purposes. The sections of the QI Tool that are relevant to the current study will be described below.

Collateral Information. The “Collateral Information” section of the QI Tool focuses on the collection of background information. Collateral interviews are defined in the QI tool as “contacts with persons or third parties who may provide relevant information to address the evaluation’s referral questions.” Collateral interviews can include interviews with resource parents, family members, therapists, medical professionals, case workers, or other individuals who may provide meaningful information about the case. Collateral records are defined in the QI tool as “documentation including reports from schools, health care providers, previous evaluations, and relevant legal documents that provide pertinent information to address the evaluation’s referral question.” Collateral data is used to gather an objective history to substantiate findings from the present evaluation. This section ends with a question asking whether or not sufficient data were included. It is up to the peer reviewer’s judgment to determine if there was enough information to answer the referral questions. If there was not, the peer reviewer is asked to explain why they believe there is not enough information to answer the referral question.

Variable: Total Background Information. A variable called “Total Background Information” (TBI) was computed that is a sum of the number of collateral sources used in each

of the aforementioned categories. This variable was an independent variable for question one and a dependent variable for question two. For the present study, the number of sources of background information was used as a proxy for the number of hours spent on an evaluation, as specified in some aforementioned sources of literature (Ackerman & Ackerman, 1996; Ackerman & Brey Pritzl, 2011; Keilin & Bloom, 1986; LaFortune & Carpenter, 1998).

Psychological Inventories and Interpretation. The “Psychological Inventories and Interpretation” section of the QI Tool includes items about psychological testing procedures. Peer reviewers are asked to identify all psychological tests used in the evaluation (excluding neuropsychological tests), whether the purpose of the tests was adequately described, and whether the provider described how to interpret the results of the tests. Commonly used psychological tests were divided into four categories in this section. The psychological testing categories were identified and defined as follows:

- **Cognitive/Achievement:** Cognitive assessments are standardized measures of individual ability in regard to specific areas (e.g., visual-spatial skills, working memory, and processing speed). Achievement assessments are standardized measures of acquired skill or knowledge (e.g., arithmetic, reading comprehension, and writing).
- **Multiscale:** Broad screening measures for a variety of personality traits and/or clinical symptoms. Elevated subscales of these measures can be further explored with specific clinical scales.
- **Clinical/Personality:** Measures designed to assess the severity of specific symptoms (e.g., depression, anxiety, and anger) and the presence of certain personality traits.

- **Child-Related Forensic Issues:** Forensically relevant assessments are typically selected based on the referral question of an evaluation. These can include measures of parenting skills, traumatic stress, substance use, or other risk assessments (Ocasio, Forsythe, Diaz, & Springer, 2018).

Additionally, there is a section that explores whether there were threats to validity in relation to culture, examinee characteristics, test administration procedures, or elevated validity indices.

Variable: Total Tests. A variable called “Total Tests” (TT) was computed that is a sum of the number of tests used in each of the aforementioned categories. This variable was used as an independent variable for question three and a dependent variable for question four.

Evaluation Rubric. The final permanent section of the QI Tool is the “Evaluation Rubric.” This section is comprised of items that assess the degree to which the evaluation meets certain criteria based on the standards outlined in the “New Jersey Department of Children and Families Guidelines for Expert Evaluations in Child Abuse/Neglect Proceedings (Mental Health)” (2012). The questions in this section are measured on a six-point Likert-type scale with “1” being “Strongly Disagree” and “6” being “Strongly Agree.” Because of the lack of quality improvement information available for forensic evaluations in child welfare (Combalbert, Andronikof, Armand, Robin, & Bazex, 2014; Grisso, 2010; Wettstein, 2005), this aspect of the NJCC study seeks to measure aspects of quality to broadly improve forensic evaluations in child welfare.

Variable: Use of a multimodal approach. This item of the QI Tool assesses the degree to which the evaluator employed a multimodal (i.e., use of multiple sources) approach in the evaluation that is being reviewed. This variable was the independent variable for question two. While the majority of available literature on forensic assessments discusses whether the

evaluation used information from multiple sources, no studies have assessed the number of unique sources from which information was gathered.

Variable: *Expertise in testing.* This item of the QI Tool assesses the degree to which the evaluator displayed expertise with the psychological tests used in the evaluation. In this case, expertise refers to the evaluator adequately describing the purpose of each test they used, describing how the tests are interpreted, using the most recent version of the test (when appropriate), and describing any threats to the validity of the test administration. This variable was the independent variable for question four.

Variable: *Overall quality.* The penultimate item on the QI Tool asks the peer reviewer to rate the evaluation's overall quality based on their answers to the prior questions in this section. This variable serves as the dependent variable for questions one and three. Although previous literature discusses the quality of forensic assessments (Fuger et al., 2014; Hecker & Steinberg, 2002; Lander & Heilbrun, 2009; Nguyen et al., 2011; Petrella & Poythress, 1983; Sanschagrin, 2006), there is still a lack of information about how to quantitatively measure quality.

Sample

A total of 1772 peer reviews were conducted for NJCC. Of these, 129 were conducted to assess reliability of the QI Tool and are excluded from this analysis. One review was omitted from the study because it did not have a final rating of overall quality, leaving a sample size of 1642.

Analytic Approach

Descriptive statistics of the total number of evaluations reviewed across the catchment areas and the mean rating for the multimodal approach, psychological test expertise, and overall quality questions were calculated. The first research question used Spearman's correlation to

determine whether there is a relationship between the number of unique sources of background information used in a forensic evaluation in child welfare and the overall quality of the evaluation. The data were then assessed for multicollinearity and because the variables were not highly multicollinear, an ordinal regression analysis was conducted to determine how the measure of overall quality can be predicted by the number of unique data sources used. The second research question used a test of ranks to examine the degree to which a multimodal approach was used to draw conclusions about the case. The test of ranks used was dependent on the distribution of the data, which was assessed using a visual inspection of the data.

Descriptive statistics were also calculated for the third research question to determine the frequency of psychological tests that are used in forensic evaluations in child welfare. Similarly to the first question, Spearman's correlation was used to determine whether there is a relationship between the number of psychological tests used and the overall quality of the evaluation. After the data were determined not to be highly multicollinear, an ordinal regression analysis was conducted to determine how the measure of overall quality can be predicted by the number of psychological tests used. The fourth research question used a test of ranks to determine if the number of psychological tests used in an evaluation differs based on the evaluator's demonstrated level of expertise using and interpreting the tests.

Results

Data Preparation

Prior to analysis, the data were screened for missing values using IBM SPSS Statistics. The total number of cases in the sample was 1643. For questions one and two, a missing values analysis indicated .01% of the sample contained missing data (n=12). Further analysis of these samples indicated that they were missing a score on the measure of overall quality and/or the

measure of whether a multimodal approach was used by the evaluator. Listwise case deletion was used to exclude any cases with missing values. As such, the total number of cases included in the analyses for questions one and two is 1631. The number of collateral interviews conducted was added to the number of unique sources of background information, resulting in a sum referred to as TBI.

For questions three and four, a missing values analysis indicated less than .01% of the sample contained missing data (n=1). This single data point was missing a score on the measure of overall quality. Listwise case deletion was used to exclude this value, leaving a sample of 1642 cases. The number of psychological tests used in each evaluation was summed into one variable called “Total Tests” (TT).

Analyses for Questions 1 and 2

Descriptive statistics of the sample are presented in Table 1. Of the 1631 cases, 92.6% (n=1511) included a review of background information. Forty-seven and six tenths percent (n=760) accessed collateral interviews and 88.7% (n=1446) accessed third party records. The mean number of collateral interviews accessed is .645 with a standard deviation of .829. The mean number of records accessed is 2.158 with a standard deviation of 1.631. The mean of TBI is 2.793 with a standard deviation of 1.944. Regarding the age of the subject of the evaluation, 29.0% of evaluations had a child as the subject (n=473) and 71.0% had an adult as the subject (n=1158).

Table 1

Descriptive Statistics for Questions 1 and 2

Background Info Used	Mean	Std. Deviation	Range (Min-Max)
Records Reviewed	2.158	1.631	0-11
Collateral Interviews	0.635	0.829	0-6
TBI	2.793	1.944	0-14

Table 2

Descriptive Statistics for Overall Quality for Question 1

Overall Quality	<i>N</i>	Percent	Cumulative Percent
1	29	1.8	1.8
2	245	15.0	16.8
3	348	21.3	38.1
4	485	29.7	67.9
5	444	27.2	95.1
6	80	4.9	100.0

Question 1. Descriptive statistics indicating the overall quality of the evaluations are presented in Table 2. The mean score for overall quality is 3.81 with a median of 4.00. A Spearman's rank-order correlation was run to assess the relationship between TBI and the overall quality of the evaluation. Preliminary analyses showed the relationship to be monotonic as assessed by visual inspection of a scatterplot. The results of the Spearman's correlation indicated that there is a statistically significant, weak positive correlation between TBI and the overall quality of the evaluation ($r_s(1629) = .260, p < .001$). Therefore, we can reject the null hypothesis and accept the alternative hypothesis.

When stratified by the age of the subject with respect to children, the results of the Spearman's rank-order correlation indicated that there is a statistically significant, very weak positive correlation between TBI and overall quality ($r_s(471) = .194, p < .001$). When stratified by the age of the subject with respect to adults, the results of the Spearman's rank-order correlation indicated that there is a statistically significant, weak positive correlation between TBI and overall quality ($r_s(1156) = .261, p < .001$).

Because there is a statistically significant relationship between the variables of interest, an ordinal logistic regression analysis was conducted to evaluate whether the measure of overall quality could be predicted by TBI and the age of the subject of the evaluation. First, the data

were assessed for multicollinearity and it was determined that the variables are not multicollinear, meeting this key assumption of ordinal logistic regression. Next, a series of separate binomial logistic regressions were run on cumulative dichotomous dependent variables to test for the assumption of proportional odds. In variables with large sample sizes, the full likelihood ratio test that compares the fit of the proportional odds model to a model with varying location parameters can lead to statistically significant results (i.e., failure of the proportional odds assumption) even when there are very small differences in slopes (Garson, 2014). Thus, further analyses were conducted to examine the estimated parameters for each binomial logistic regression run on each dichotomized cumulative category. After examining the estimated parameters, it was determined that the model did not meet the assumption of proportional odds.

When examining the output of this ordinal regression analysis, a warning was generated by SPSS stating that there were 29.3% of cells with zero frequency. Consensus in the field of statistical analysis dictates that there should be 80% or more expected cell frequencies to reliably interpret the goodness-of-fit measures (Garson, 2014). Thus, to achieve adequate cell frequencies, the categories of the ordinal dependent variable were collapsed. Upon collapsing the categories of the ordinal dependent variable from six to three, the percentage of cells with zero frequency decreased from 29.3% to 16.0%. Additionally, the assumption of proportional odds became tenable ($\chi^2(2) = 4.738, p = .094$).

The deviance goodness-of-fit test indicated that the model was a good fit for the observed data, $\chi^2(46) = 55.497, p = .132$. Table 3 depicts the results of the regression analysis. The final model statistically significantly predicted the dependent variable over and above the intercept-only model, $\chi^2(2) = 101.553, p < .001$. The odds (OR) of an evaluator using more background information sources on an evaluation with a child as the subject is 1.247, 95% CI [1.011, 1.538],

a statistically significant effect, $\chi^2(1) = 4.252, p = .039$. An increase TBI was associated with an increase in the odds of having a higher rating of overall quality with an odds ratio of 1.267, 95% CI [1.204, 1.334], Wald $\chi^2(1) = 81.567, p < .001$.

Table 3

Summary of Logistic Regression Analysis for TBI and Age Predicting Overall Quality

Source	B	SE B	Wald χ^2	p	OR	95% CI OR
Subject - Child	.221	.107	4.252	.039*	1.247	[1.011, 1.538]
TBI	.237	.026	81.567	.000**	12.67	[1.204, 1.334]

* $p < .05$, ** $p < .001$

Question 2. Descriptive statistics indicating the degree to which a multimodal approach was used to draw conclusions about the case are presented in Table 4. The mean score for this independent variable is 3.98 with a median of 4.00. This variable was collapsed into three categories of “low,” “medium,” and “high.” Before conducting the test of ranks, the data were assessed for the assumption of normality using the visual inspection of a boxplot and histogram. The data met this assumption, though there were outliers present, so it was determined that the non-parametric Kruskal-Wallis test would be appropriate.

Table 4

Descriptive Statistics for Multimodal Approach

Multimodal Approach Used	N	Percent	Cumulative Percent
1	26	1.6	1.6
2	193	11.8	13.4
3	322	19.7	33.2
4	503	30.8	64.0
5	422	25.9	89.9
6	165	10.1	100.0

A Kruskal-Wallis test was conducted to determine if there were differences in the number of TBI used between groups that differed on the measure of whether the evaluator used a multimodal approach to draw conclusions about the case: the “low” ($N = 219$), “medium” ($N =$

825), and “high” ($N = 587$) groups. Distributions of the number of scores were similar for all groups as assessed by visual inspection of a boxplot. Median scores were statistically significant between groups, $\chi^2(2) = 200.360, p = <.001$. Pairwise comparisons were performed using Dunn’s (1964) procedure with a Bonferroni correction for multiple comparisons. This post hoc analysis revealed statistically significant differences in median number of TBI between all combinations of the low, medium, and high groups ($p < .001$).

Analyses for Questions 3 and 4

Descriptive statistics of the sample are presented in Table 5. Of the 1642 cases, 90.9% ($N=1493$) used a psychological test as part of the evaluation. The mean number of tests used in each evaluation is 3.01 with a standard deviation of 1.848.

Table 5

Descriptive Statistics for Questions 3 and 4

Type of Test Used	Mean	Std. Deviation	Range (Min-Max)
Forensic	.824	.906	0-6
Clinical	.6768	.919	0-4
Multiscale	.9251	.633	0-4
Cognitive	.589	.787	0-6
TT	3.014	1.848	0-11

Table 6

Descriptive Statistics for Overall Quality for Question 3

Overall Quality	N	Percent	Cumulative Percent
1	30	1.8	1.8
2	245	14.9	16.7
3	349	21.2	38.0
4	489	29.8	67.8
5	448	27.3	95.1
6	81	4.9	100.0

Question 3. Descriptive statistics indicating the overall quality of the evaluations are presented in Table 6. As previously stated, the mean score for overall quality is 3.81 with a

median of 4.00. A Spearman's rank-order correlation was run to assess the relationship between TT and the overall quality of the evaluation. Preliminary analyses showed the relationship to be monotonic as assessed by visual inspection of a scatterplot. The results of the Spearman's correlation indicated that there is a statistically significant, very weak positive correlation between the total background information used in a forensic evaluation in child welfare and the overall quality of the evaluation ($r_s(1640) = .086, p < .001$). Therefore, we can reject the null hypothesis and accept the alternative hypothesis.

When stratified by the age of the subject with respect to children, the results of the Spearman's rank-order correlation indicated that there is a statistically significant, very weak positive correlation between TT and overall quality ($r_s(475) = .136, p < .001$). When stratified by the age of the subject with respect to adults, the results of the Spearman's rank-order correlation indicated that there is a statistically significant, very weak positive correlation between TT and overall quality ($r_s(1163) = .134, p < .001$).

Because there is a statistically significant relationship between the variables of interest, an ordinal logistic regression analysis was conducted to evaluate whether the measure of overall quality could be predicted by TT and the age of the subject of the evaluation. After it was determined that the data is not multicollinear, a series of separate binomial logistic regressions were run on cumulative dichotomous dependent variables to test for the assumption of proportional odds. After examining the estimated parameters, it was determined that the model did not meet the assumption of proportional odds.

When examining the output of this ordinal regression analysis, a warning was generated by SPSS stating that there were 21.9% of cells with zero frequency. Thus, to achieve adequate cell frequencies, the categories of the ordinal dependent variable were collapsed. Upon

collapsing the categories of the ordinal dependent variable from six to three, the percentage of cells with zero frequency decreased from 21.9% to 15.2%. Additionally, the assumption of proportional odds became tenable based on the results of the estimated parameters of the binomial logistic regressions.

The deviance goodness-of-fit test indicated that the model was not a good fit for the observed data, $\chi^2(40) = 90.467, p < .001$; however, the final model statistically significantly predicted the dependent variable over and above the intercept-only model, $\chi^2(2) = 33.764, p < .001$. Table 6 depicts the results of the regression analysis. The odds of an evaluator using more psychological tests on an evaluation with a child as the subject is 1.756, 95% CI [1.413, 2.183], a statistically significant effect, $\chi^2(1) = 25.710, p < .001$. An increase in the number of TT used was associated with an increase in the odds of having a higher rating of overall quality with an odds ratio of 1.125, 95% CI [1.066, 1.187], Wald $\chi^2(1) = 18.673, p < .001$.

Table 7

Summary of Logistic Regression Analysis for TT and Age Predicting Overall Quality

Source	B	SE B	Wald χ^2	p	OR	95% CI OR
Subject - Child	.563	.111	95.794	.000**	1.756	[1.413, 2.183]
TT	.118	.027	18.673	.000**	1.125	{1.066, 1.187}

* $p < .05$, ** $p < .001$

Question 4. Descriptive statistics indicating the degree to which the evaluator showed expertise with the instruments employed in the evaluation are presented in Table 8. Of the 1642 evaluations in the total sample, 1491 included a rating for this measure. The mean score for this independent variable is 4.03 with a mediation of 4.00. This variable was collapsed into three categories of “low,” “medium,” and “high.” Before conducting the test of ranks, the data were assessed for the assumption of normality using the visual inspection of a boxplot and histogram.

The data met this assumption, though there were outliers present, so it was determined that the non-parametric Kruskal-Wallis test would be appropriate.

Table 8

Descriptive Statistics for Level of Expertise

Level of Expertise	<i>N</i>	Valid Percent	Valid Cumulative Percent
1	72	4.8	4.8
2	141	9.5	14.3
3	260	17.4	31.7
4	389	26.0	57.8
5	450	30.2	88.0
6	179	12.0	100.0

A Kruskal-Wallis test was conducted to determine if there were differences in the number of psychological tests used between groups that differed on the measure of whether the evaluator used demonstrated expertise in using psychological tests: the “low” ($N = 213$), “medium” ($N = 649$), and “high” ($N = 629$) groups. Distributions of the number of scores were similar for all groups as assessed by visual inspection of a boxplot. Median scores were statistically significant between groups, $\chi^2(2) = 39.341$, $p = <.001$. Pairwise comparisons were performed using Dunn’s (1964) procedure with a Bonferroni correction for multiple comparisons. This post hoc analysis revealed statistically significant differences in median number of TT between all combinations of the low, medium, and high groups ($p < .001$).

Discussion

Producing high quality forensic evaluations in child welfare is extremely important as the findings of these evaluations can lead to life-changing events for children and families. As such, the present study sought to fill a gap in the literature on child welfare by examining the relationship between the content of forensic evaluations in child welfare and the quality of these evaluations with respect to the use of background information and psychological testing

procedures. There is a body of existing literature that addresses the content of various other forms of forensic assessment (e.g., child custody evaluations, adult criminal evaluations, and juvenile evaluations), although literature that specifically addresses forensic evaluations in child welfare is scarce. Additionally, there is a lack of quality improvement research available for forensic assessments, especially ones involving child welfare matters (Combalbert et al., 2014, Grisso, 2010; Wettstein, 2005).

One previous study used a quality improvement tool called the STAR to provide objective ratings of quality, though it did not discuss whether certain features of reports led to higher quality reports (Duits, van der Hoorn, Wiznitzer, Wettstein, & de Beurs, 2012). A series of studies (Fuger, 2014; Lander & Heilbrun, 2009, Nguyen et al., 2011; Sanschagrin, 2006) used quality improvement tools based on Heilbrun's (2001) principles of forensic mental health assessment examining criminal forensic evaluations. The present study is the first study known to this author that examines the content of forensic evaluations in child welfare as predictors of overall quality.

Previous literature used survey data to report on the percentage of practitioners that used each source of data (e.g. records, collateral interviews, and specific psychological tests) and the number of hours spent conducting each aspect of the evaluation; however, the present study is unique in that it examines the total number of unique background sources and number of psychological tests used in each report. To glean additional information from the data, the relationships were also examined when factoring in the variable of the age of the subject of the evaluation. The information used in this study was gathered through the QI Tool, a measure designed by NJCC, to assess the quality of forensic evaluations in child welfare with the ultimate goal of making recommendations to improve the quality and utilities of these evaluations. The

findings of this study suggest that there is a relationship between the amount of data used as part of a forensic evaluation in child welfare and its overall quality.

Background Information and Quality

The first two questions of this study relate to the use of background information in forensic evaluations in child welfare. The first question specifically examined the relationship between the number of unique sources of background information used in a forensic evaluation in child welfare and the overall quality of the evaluation as measured on a Likert-type scale. Additionally, this question sought to determine if the overall quality score can be predicted by the number of background sources used. The second question looked at the measure of whether a multimodal approach (i.e., multiple sources) was used for the evaluation and if the number of unique sources of background information differed based on that metric.

Descriptive statistics for questions one and two were consistent with previous literature in that the majority of forensic evaluations in child welfare use background information to draw conclusions about the case (Ackerman & Ackerman, 1996; Bow & Quinnell, 2002; Heilbrun et al., 1994; LaFortune & Carpenter, 1998; Neal & Grisso, 2014; Quinnell & Bow, 2001). None of the available literature indicated the number of unique background sources used within the forensic evaluation; however, four papers indicated the amount of time spent completing certain activities related to conducting forensic evaluations (Ackerman & Ackerman, 1996; Ackerman & Brey Pritzl, 2011; Keilin & Bloom, 1986; LaFortune & Carpenter, 1998), thus, the total number of unique background sources in an evaluation was used as a proxy for the time spent conducting record reviews and collateral interviews.

The results of question one indicated that the number of unique background sources used in an evaluation was significantly, but weak positive correlation with the overall quality of the

evaluation. To the author's knowledge, no previous studies have examined the quality of evaluations using this metric. However, the available literature shows that the amount of time spent on forensic evaluations has substantially increased since Keilin and Bloom (1986) conducted their survey of child custody practices (Ackerman & Brey Pritzl, 2011), and in viewing the number of unique background sources as a proxy, it can be argued that forensic evaluators are gathering more sources of third-party information for their reports.

When stratified by the age of the subject with respect to adults, the number of background sources remained statistically significant with a weak positive correlation with the overall quality of the evaluation. When stratified by the age of the subject with respect to children, the number of background sources remained statistically significant, yet the strength of the correlation decreased from weak to very weak when correlated with the measure of overall quality. An ordinal logistic regression analysis was run with the predictors of age and number of unique sources of background information. The results of this regression analysis indicated that an increase in the number of unique sources of background information used in a forensic evaluation in child welfare is significantly associated with an increase in the score of overall quality. Additionally, the evaluators were more likely to use more unique sources when a child was the subject of the evaluation. This result is consistent with the findings of Petrella and Poythress (1983) who found that reports that used multiple, external sources of information were more thorough and thus, of higher quality.

The results of question two showed that there are significant differences in the median number of background sources used between groups that differed on the measure of whether the evaluator used a multimodal approach to draw conclusions about the case. While there is no previous literature that specifically examines this construct, it can be argued that the findings of

Petrella and Poythress (1983) also apply here as evaluations that were deemed to be of higher quality included multiple collateral sources. Additionally, Lander and Heilbrun (2009) found that evaluations that adhered to more of Heilbrun's (2001) outlined principles were deemed to be higher quality. The principles related to the use of background information include: Use Multiple Sources of Information For Each Area Being Assessed, Obtain Relevant Historical Information, and Use Third Party Information in Assessing Response Style (Lander & Heilbrun, 2009). Thus, the results of the current study were consistent with Lander and Heilbrun's (2009) findings that using more sources of information leads to a higher quality evaluation. However, it is important to note that an evaluation should not be deemed of lower quality simply because it lacks a sufficient amount of third-party information. Heilbrun et al. (1994) noted that evaluators may struggle obtaining information from other treatment providers and it is best practice to indicate when sources were requested, but were ultimately unavailable for review (Otto et al., 2014).

Psychological Testing and Quality

Questions three and four are similar to questions one and two, respectively, though they examined psychological test usage and whether the evaluator demonstrated expertise in psychological testing procedures. Specifically, question three examined the relationship between the number of psychological tests used in a forensic evaluation in child welfare and the overall quality of the evaluation. It also examined whether the overall quality of the evaluation can also be predicted by the number of psychological tests used. Question four looks at the measure of whether the evaluator demonstrated expertise in their use of psychological tests and if the number of psychological tests used differed based on that level of demonstrated expertise. Previous literature shows that psychological testing plays a significant role in the process of conducting forensic assessments (Ackerman & Ackerman, 1996; Ackerman & Brey Pritzl, 2011;

Hagen & Castagna, 2001; Keilin & Bloom, 1986; Lally, 2003; Neal & Grisso, 2014; Quinnell & Bow, 2001). Generally, psychological testing in forensic assessment is seen as another way to obtain data without relying solely on examinee interviews (Brodzinsky, 1993; Neal & Grisso, 2014).

The results of this study showed that psychologists are using slightly fewer psychological tests in forensic evaluations for child welfare than reported in previous literature (Ackerman & Ackerman, 1996; Neal & Grisso, 2014). The descriptive statistics showed that forensic practitioners, on average, use less than one test per defined category. Previous NJCC research showed that providers in New Jersey most often rely on multiscale inventories for the majority of evaluations (Ocasio et al., 2018), although the descriptive statistics from the present study showed that not all providers consistently use these types of measures.

Similarly to the results of question one, the number of psychological tests used in an evaluation are both positively, but weakly, associated with the overall quality of the evaluation. To the author's knowledge, no previous studies have examined the quality of evaluations using this metric. When stratified by the age of the subject, the number of psychological tests used remained statistically significant, yet very weakly correlated with the measure of overall quality. An ordinal logistic regression analysis was run with the predictors of age and number of psychological tests. The results of this regression analysis indicated that an increase in the number of psychological tests used in a forensic evaluation in child welfare is significantly associated with an increase in the score of overall quality. Additionally, the evaluators were more likely to use more psychological tests when a child was the subject of the evaluation. It is possible that more psychological tests are used when children are the subject of the evaluation due to a need for more data when assessing children, as clinical interviews with younger children

may not be as reliable as those with adults. There is no extant literature that specifically describes the relationship between psychological tests and quality of an evaluation, though it can be argued that these results are consistent with Lander and Heilbrun's (2009) work in which higher quality reports adhered more closely to the *Principles of Forensic Mental Health Assessment* (Heilbrun, 2001). The relevant principles include: Use Multiple Sources of Information For Each Area Being Assessed, Use Testing When Indicated in Response Style, and Use Nomothetic Evidence in Assessing Clinical Condition, Functional Abilities and Causal Connection (Heilbrun, 2001).

The results of question four showed that there are significant differences in the median number of psychological tests used between groups that differed on the measure of whether the evaluator demonstrated expertise in using the tests including how the results were reported and whether there were any concerns about the validity of the test administration. There is no specific literature that discusses this construct, so the information gathered from this aspect of the study can inform future studies. Ultimately, the findings of this paper are strongly in favor of using multiple sources of data collection to inform forensic evaluations in child welfare, including background information and psychological testing results.

Limitations of the Study

Results from this study are exploratory, as it is the first to examine the relationship between the content and quality of forensic evaluations in child welfare with respect to background information and psychological testing. Although the production of high quality reports is important in all areas of forensic and psychological assessment, the generalizability of this study's findings may be limited based on the psycholegal and referrals questions being asked as part of the evaluation. The study also only includes data from one state and New Jersey's procedures for conducting these evaluations may differ from other jurisdictions. Future research

may benefit from examining samples collected nationally to have a clearer idea of professional practice throughout the United States.

The data for this study was gathered solely through the QI Tool. The QI Tool has good reliability (Koo & Ti, 2016), although there are some flaws in its design. For example, the Likert-type items are measured on a six-point scale; as such, there is no neutral or middle option to measure the quality of the evaluation (Froman, 2014). The Likert-type items are also labeled as “Strongly Disagree” for a rating of “1” and “Strongly Agree” for a rating of “6,” though these verbal anchors did not always apply to the item. Rater bias is also a limitation of this study due to the single method of data collection. Finally, although the reliability of the QI Tool is good, it is possible that if it was higher, the strengths of the correlations would increase. The QI Tool has increased in reliability across versions (Version 2 ICC = .680 versus Version 3 ICC = .824), which could also impact the results, as the data were collected on different versions of the QI Tool. Despite these flaws, the development and utilization of the QI Tool can serve as a model for quality improvement in forensic evaluations in child welfare around the United States.

Regarding the analytical processes, while the relationships examined by the data are statistically significant, they were weakly correlated. Replication of this study could include more in-depth analyses of these correlations in addition to examining different stratification variables, rather than the age of the subject of the evaluation. Additionally, this study utilized the proportional odds model to model the dependence of the ordinal measure of overall quality on the continuous variables of the number of unique sources of background information and the number of psychological tests used in an evaluation. This model relies on the data meeting the assumption of proportional odds, which the data violated before the categories were collapsed. Depending on the data, collapsing the categories in an ordinal regression analysis can affect the

effect estimate (Strömberg, 1996). A partial proportional odds model could be conducted on the uncollapsed data as it does not require that the data meets the assumption of proportional odds (Peterson & Harrell, 1990). The author was unable to utilize the partial proportional odds model due to limitations in the availability of certain statistical computing software. Future research may also use the Jonckheere-Terpstra test for questions two and four in lieu of the Kruskal-Wallis test to determine if there is a statistically significant monotonic trend between an ordinal independent variable and a continuous dependent variable. The Jonckheere-Terpstra test uses the ordinal nature of the independent variable to test for trends, thus producing more meaningful results.

This study did not examine whether specified unique sources of background information or psychological tests produced higher quality reports. Future research would benefit from determining whether it is more or less valuable to gain information from third-party sources or administer certain psychological tests depending on a variety of factors, including the referral questions sought to be answered. Additionally, it may be beneficial to investigate whether there is an “ideal” number of sources of background information or psychological tests that should be used.

Another limitation of this study is that the peer reviewers for NJCC are all individuals with some degree of training in forensic psychology. Future research could examine whether the consumers of the reports (e.g., lawyers, judges, and DCP&P workers) agree with the ratings of quality and, if not, how their ratings compare to those of the peer reviewers. Additionally, the reviews are not conducted blindly. This means that peer reviewers are aware of which practitioners are writing the evaluations. This is not to say that peer reviewers would

purposefully assign higher or lower ratings to certain practitioners, however, it is possible that peer reviewers may show bias toward or against other practitioners with whom they are familiar.

Implications for Research and Practice

Forensic practitioners often view their work as objective, though there is research that suggests this may not be the case (Neal & Brodsky, 2016; Zapf & Dror, 2017). Thus, Lally's (2003) suggestion for the inclusion of psychological testing data and background information, in addition to a clinical interview of the subject of the evaluation, is especially important. Results from this study can add to the limited existing literature on the content of forensic evaluations in child welfare; additionally, these results can serve as a model for future studies that examine the quality of forensic assessments, specifically on those conducted for child welfare or child protection matters. While there is evidence that using more background sources and/or more psychological tests produces higher quality reports, practitioners should still be conscious that they are answering the referral question using appropriate background sources and psychological tests, rather than using these sources of information to boost the perceived quality of the report.

Several researchers have identified areas for improvement within the field of forensic psychology (Grisso, 2010; Neal & Grisso, 2014; Wettstein, 2005), though there is still limited research available on the utility of these suggestions. While the APA has various specialty guidelines relating to forensic practice (2010; 2013a, 2013b), they are only guidelines and no psychologist is required to adhere to the outlined recommendations. There remains no established standards of practice or care for forensic evaluations in child welfare. Peer review can serve as an important means of improving the quality and utility of these evaluations (Neal &

Brodsky, 2016), thus the data from this study and NJCC overall can ultimately play a major role in improving outcomes for children and families who are involved in the child welfare system.

One area of interest for the author relates to how this study might impact the field of school psychology. At surface level, it may seem that forensic psychology and school psychology do not have much in common; however, much like forensic practitioners, school psychologists rely heavily on third-party and objective testing data to draw conclusions. School psychologists use testing data to classify students into special education categories, which ultimately has a major impact on a student's educational trajectory. Psychoeducational evaluations can be subject to legal scrutiny as well, so it is imperative that high quality reports are produced with plenty of evidence to support the findings. **Conclusion**

Using multiple sources of information to draw conclusions is imperative to producing high quality forensic evaluations in child welfare. The results of the current study found that the more unique sources of background information used to inform a forensic evaluation in child welfare, the higher it scored on measure of overall quality. Similar results were found when examining the use of psychological tests in forensic evaluations in child welfare. Forensic practitioners should strive to integrate multiple sources of data into their evaluations, as producing high quality evaluations can ultimately lead to improved outcomes for families who are involved in the child welfare system. Future research may benefit from examining procedures and standards for conducting forensic evaluations in child welfare in other jurisdictions as well as more specific exploration of the types of sources that tend to improve the quality of forensic evaluations in child welfare.

References

- Ackerman, M. J., & Ackerman, M. C. (1996). Child Custody Evaluation Practices: A 1996 Survey of Psychologists. *Family Law Quarterly*, *30*(3), 565-586.
- Ackerman, M. J., & Brey Pritzl, T. (2011). Child custody evaluation practices: A 20-year follow-up. *Family Court Review*, *49*(3), 618–628. doi:10.1111/j.1744-1617.2011.01397.x
- Aimola, L., Jasim, S., Tripathi, N., Tucker, S., Worrall, A., Quirk, A., & Crawford, M. J. (2016). Impact of peer-led quality improvement networks on quality of inpatient mental health care: study protocol for a cluster randomized controlled trial. *BMC Psychiatry*, *16*, 331. <http://doi.org/10.1186/s12888-016-1040-1>
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing*. Washington, DC: AERA
- American Psychological Association (2009) Criteria for the Evaluation of Quality Improvement Programs and the Use of Quality Improvement Data. (2009). *American Psychologist*, *64*(6), 551–557. <https://doi.org/10.1037/a0016744>
- American Psychological Association (2010). Guidelines for Child Custody Evaluations in Family Law Proceedings. *American Psychologist*, *65*(9), 863-867.
- American Psychological Association (2013). Guidelines for Psychological Evaluations in Child Protection Matters. *American Psychologist*, *68*(1), 20-31.
- American Psychological Association. (2013). Specialty Guidelines for Forensic Psychology. *American Psychologist*, *68*(1), 377-402.

- Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment, 87*, 84–94.
- Bow, J., & Quinnell, F. (2002). A critical review of child custody evaluation reports. *Family Court Review, 40*(2), 164–176. <https://doi.org/10.1111/j.174-1617.2002.tb00827.x>
- Bricklin, B. (1984). *Bricklin Perceptual Scales Manual*. Furlong, PA: Village Publishing.
- Brydon-Miller, M. (1997). Participatory action research: Psychology and social change. *Journal of Social Issues, 53*(4), 657-666.
- Cates, J. (1999). The art of assessment in psychology: Ethics, expertise, and validity. *Journal of Clinical Psychology, 55*(5), 631–641.
- Cheng, H. G., & Phillips, M. R. (2014). Secondary analysis of existing data: opportunities and implementation. *Shanghai Archives of Psychiatry, 26*(6), 371–375.
<http://doi.org/10.11919/j.issn.1002-0829.214171>
- Combalbert, N., Andronikof, A., Armand, M., Robin, C., & Bazex, H. (2014). Forensic mental health assessment in France: Recommendations for quality improvement. *International Journal of Law and Psychiatry, 37*(6), 628-634.
- Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. 2786 (1993).
- Donaldson, S., & Grant-Vallone, E. (2002). Understanding Self-Report Bias in Organizational Behavior Research. *Journal of Business and Psychology, 17*(2), 245-260.
- Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics, 6*, 241-252.
- Duits, N., van der Hoorn, S., Wiznitzer, M., Wettstein, R. M., & de Beurs, E. (2012). Quality improvement of forensic mental health evaluations and reports of youth in the Netherlands. *International Journal of Law and Psychiatry, 35*(5-6), 440-444.

- Edens, J., & Boccaccini, M. (2017). Taking Forensic Mental Health Assessment “Out of the Lab” and into “the Real World”: Introduction to the Special Issue on the Field Utility of Forensic Assessment Instruments and Procedures. *Psychological Assessment, 29*(6), 599–610. <https://doi.org/10.1037/pas0000475>
- Froman, R. (2014). The Ins and Outs of Self-Report Response Options and Scales. *Research in Nursing & Health, 37*(6), 447–451. <https://doi.org/10.1002/nur.21626>
- Fuger, K., Acklin, M., Nguyen, A., Ignacio, L., & Gowensmith, W. (2014). Quality of criminal responsibility reports submitted to the Hawaii judiciary. *International Journal of Law and Psychiatry, 37*(3), 272–280. <https://doi.org/10.1016/j.ijlp.2013.11.020>
- Garson, G. D. (2014). *Ordinal Regression*. Asheboro, NC: Statistical Associates Publishers.
- Grisso, T. (2010). Guidance for improving forensic reports: A review of common errors. *Open Access Journal of Forensic Psychology, 2*, 102–115.
- Grol, R. (1994). Quality improvement by peer review in primary care: a practical guide. *Quality in Health Care, 3*(3), 147–152.
- Hagen, M. A., Castagna, N. (2001) The real numbers: Psychological testing in custody evaluations. *Professional Psychology: Research and Practice, 32*(3), 269-271.
- Hayes, S., Nelson, R., & Jarrett, R. (1987). The Treatment Utility of Assessment. *American Psychologist, 42*(11), 963–974. <https://doi.org/10.1037/0003-066X.42.11.963>
- Hecker, T. & Steinberg, L. (2002). Psychological Evaluation at Juvenile Court Disposition. *Professional Psychology: Research and Practice, 33*(3), 300–306. <https://doi.org/10.1037/0735-7028.33.3.300>
- Heilbrun, K. (1992). The Role of Psychological Testing in Forensic Assessment. *Law and Human Behavior, 16*(3), 257–272. <https://doi.org/10.1007/BF01044769>

Heilbrun, K. (1995). Child custody evaluation: Critically assessing mental health experts and psychological tests. *Family Law Quarterly*, 29(1), 63-78.

Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York: Kluwer Academic/Plenum Publishers

Heilbrun, K., Dematteo, D., Marczyk, G., & Goldstein, A. (2008). Standards of Practice and Care in Forensic Mental Health Assessment: Legal, Professional, and Principles-Based Considerations. *Psychology, Public Policy, and Law*, 14(1), 1–26.
<https://doi.org/10.1037/1076-8971.14.1.1>

Heilbrun, K., NeMoyer, A., King, C., & Galloway, M. (2015). Using third-party information in forensic mental-health assessment: a critical review. *Court Review*, 51(1), 16–35.

Heilbrun, K., Rogers, R., & Otto, R. K. (2002). Forensic Assessment: Current Status and Future Directions. In Ogloff, J. R. P. (Ed), *Taking psychology and law into the twenty-first century* (pp. 119-146). Boston, MA: Springer US.

Heilbrun, K., Rosenfeld, B., Warren, J., Collins, S., & Heilbrun, K. (1994). The use of third-party information in forensic assessments: a two-state comparison. *The Bulletin of the American Academy of Psychiatry and the Law*, 22(3), 399–406. Retrieved from <http://search.proquest.com/docview/77715908/>

Hodges, K. (2004). Using Assessment in Everyday Practice for the Benefit of Families and Practitioners. *Professional Psychology: Research and Practice*, 35(5), 449–456.
<https://doi.org/10.1037/0735-7028.35.5.449>

Keilin, W. G., & Bloom, L. J. (1986). Child custody evaluation practices: A survey of experienced professionals. *Professional Psychology: Research and Practice*, 17(4), 338-346.

- Koo, T., & Li, M. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine, 15*(2), 155–163.
<https://doi.org/10.1016/j.jcm.2016.02.012>
- LaFortune, K., & Carpenter, B. (1998). Custody evaluations: a survey of mental health professionals. *Behavioral Sciences & the Law, 16*(2), 207–224.
- Lally, S. J. (2003). What Tests Are Acceptable for Use in Forensic Evaluations? A Survey of Experts. *Professional Psychology: Research and Practice, 34*(5), 491-498.
- Lander, T. D., & Heilbrun, K. (2009). The content and quality of forensic mental health assessment: Validation of a principles-based approach. *The International Journal of Forensic Mental Health, 8*(2), 115-121. <http://dx.doi.org/10.1080/14999010903199324>
- Laulik, S., Allam, J., & Browne, K. (2013). The Use of the Child Abuse Potential Inventory in the Assessment of Parents involved in Care Proceedings. *Child Abuse Review, 24*, 332-345. <https://doi.org/10.1002/car.2294>
- Lees-Haley, P. R. (1992). Psychodiagnostic test usage by forensic psychologists. *American Journal of Forensic Psychology, 10*(1), 25-30.
- Lubin, B., Larsen, R., Matarazzo, J., Seever, M., & Pallak, M. (1985). Psychological Test Usage Patterns in Five Professional Settings, *American Psychologist, 40*(7), 857–861.
[doi:10.1037/0003-066X.40.7.857](https://doi.org/10.1037/0003-066X.40.7.857)
- Meyer, G., Finn, S., Eyde, L., Kay, G., Moreland, K., Dies, R., ... Reed, G. (2001). Psychological Testing and Psychological Assessment. *American Psychologist, 56*(2), 128–165. <https://doi.org/10.1037/0003-066X.56.2.128>
- Milner, J. S. (1986). *The Child Abuse Potential Inventory: manual* (2nd ed.). Webster, NC: Psytec.

- Neal, T., & Brodsky, S. L. (2016). Forensic psychologists' perceptions of bias and potential correction strategies in forensic mental health evaluations. *Psychology, Public Policy, and Law*, 22(1), 58.
- Neal, T., & Grisso, T. (2014). Assessment Practices and Expert Judgment Methods in Forensic Psychology and Psychiatry: An International Snapshot. *Criminal Justice and Behavior*, 41(12), 1406–1421. doi:10.1177/0093854814548449
- Nelson-Gray, R. (2003). Treatment Utility of Psychological Assessment. *Psychological Assessment*, 15(4), 521–531. <https://doi.org/10.1037/1040-3590.15.4.521>
- New Jersey Child Welfare Data Hub. (2018). Child Protective Services Reports. Retrieved on May 6, 2019 from <https://njchilddata.rutgers.edu/portal/cps-reports>.
- New Jersey Department of Children and Families. (2012). Guidelines for Expert Evaluations in Child Abuse/Neglect Proceedings (Mental Health). Trenton, NJ.
- New Jersey Department of Children and Families (2018). Request for Qualifications for Forensic Evaluation Services by Psychologists. Trenton, NJ.
- Nguyen, A., Acklin, M., Fuger, K., Gowensmith, W., & Ignacio, L. (2011). Freedom in paradise: Quality of conditional release reports submitted to the Hawaii judiciary. *International Journal of Law and Psychiatry*, 34(5), 341–348.
- Ocasio, K., Forsythe, M., Diaz, C., & Springer, C. (2018, June). *Psychological Assessment Tools: An Exploration of Usage and Appropriateness for Child Welfare Forensic Evaluations*. Poster session presented at the Annual Colloquium of the American Professional Society on the Abuse of Children, New Orleans, LA.

- Otto, R. K., DeMier, R. L., & Boccaccini, M. T. (2014). *Forensic reports and testimony: A guide to effective communication for psychologists and psychiatrists*. Hoboken, NJ, US: John Wiley & Sons Inc.
- Peterson, B., & Harrell, F. (1990). Partial Proportional Odds Models for Ordinal Response Variables. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 39(2), 205–217. <https://doi.org/10.2307/2347760>
- Petrella, R. & Poythress, N. (1983). The quality of forensic evaluations: An interdisciplinary study. *Journal of Consulting and Clinical Psychology*, 51(1), 76–85. <https://doi.org/10.1037/0022-006X.51.1.7>
- Otto, R., Edens, J., & Barcus, E. (2000). The use of psychological testing in child custody evaluations. *Family Court Review*, 38(3), 312–340. doi:10.1111/j.174-1617.2000.tb00578.x
- Quinnell, F. A., & Bow, J. N. (2001). Psychological tests used in child custody evaluations. *Behavioral Sciences and the Law*, 19, 491–501
- Robey, A. (1965). Criteria for competency to stand trial: A checklist for psychiatrists. *American Journal of Psychiatry*, 122, 616-622.
- Rogers, R., Salekin, R. T., & Sewell, K. W. (1999). Validation of the Millon Clinical Multiaxial Scale for Axis II disorders: Does it meet the Daubert standard? *Law and Human Behavior*, 23, 425-443.
- Rutgers School of Social Work Institute for Families. (n.d.). New Jersey Coordination Center for Child Abuse and Neglect Forensic Evaluation and Treatment. Retrieved from <https://socialwork.rutgers.edu/centers/institute-families/new-jersey-coordination-center-child-abuse-and-neglect-forensic-evaluation-and-treatment>

Sanschagrin, K., & Heilbrun, K. (2005). *The quality of forensic mental health assessments of juvenile offenders: An empirical investigation* (ProQuest Dissertations Publishing).

Retrieved from <http://search.proquest.com/docview/304992605/>

Stahl, P. (1999). A Child Custody Evaluation: A process designed to help your kids thrive.

Family Advocate, 22(1), 34–37.

Strömberg, U. (1996). Collapsing ordered outcome categories: a note of concern. *American*

Journal of Epidemiology, 144(4), 421–424.

<https://doi.org/10.1093/oxfordjournals.aje.a008944>

Wettstein, R. M. (2005). Quality and quality improvement in forensic mental health evaluations.

Journal of the American Academy of Psychiatry and the Law Online, 33(2), 158-175.

Zapf, P. A., & Dror, I. E. (2017). Understanding and mitigating bias in forensic evaluation:

Lessons from forensic science. *International Journal of Forensic Mental Health*, 16(3), 227-238.

Appendix A

NJCC Quality Improvement Tool

Forensic Evaluation Quality Improvement Tool: Code Book

New Jersey Forensic Coordination Center
Rutgers University School of Social Work

Revised 5-01-2018



Table of Contents

General Information Regarding the Forensic Evaluation QA Tool.....	59
Block I: Evaluator & Evaluation Detail Sheet	61
Block II: Referral Information.....	62
Block III: Case Detail Sheet.....	64
Block IV: Background & Demographic Information	66
Block V: Interview with a Child	69
Block VI: Interview with an Adult.....	72
Block VII: Psychological Inventories and Interpretation	76
Block VIII: Evaluation Questions	80
Block IX: Collateral Information	82
Block X: Bonding and Attachment	84
Block XI: Summary	85
Block XII: Recommendations	86
Block XIII: Evaluation Rubric.....	89
Glossary of Terms.....	91
References	97

General Information Regarding the Forensic Evaluation QA Tool

Response Format

- Indicating “Yes” or “No” to specific questions
 - Yes – If the practice is described as part of the evaluation
 - No – If the practice is NOT described as part of the evaluation
- N/A – Not applicable/specified for the current evaluation
 - Will only be available for a select number of questions
- Reviewer Section Comments
 - Located at the end of each block
 - Comments pertaining to the current evaluation and / or how the peer reviewer responded to a specific question while using the QA Tool
 - Responses and comments should be limited to specific questions found in the block for each “Reviewer Section Comments” text box
- Selecting a Response
 - If the QA Tool Question has a circular selection box you can only select one option.
 - If the QA Tool Question has a square selection box you may select as many options as appropriate.
 - All questions must be answered; if for any reason you feel a response is not appropriate please answer it and you may explain the selection in the Reviewer Section Comments located at the end of the block.
- As Indicated by the Evaluator
 - Certain questions specify that responses should be limited to what was stated by the Evaluator for the current evaluation - “as indicated by the Evaluator.”
 - If for any reason you disagree with what the Evaluator specified in the evaluation indicate so in the “Reviewer Section Comments” at the end of the block.

Question Coding

- B: Background
 - This question is gathering background information related to the Evaluator, forensic evaluation, or subject of the Assessment.
 - Data collected from this category is not considered when assessing the quality of the forensic evaluation.
- E: Exploratory
 - This question is gathering information about a practice or set of practices which have been observed in forensic evaluations but are not mandated by

the “Guidelines for Expert Evaluations in Child Abuse/Neglect Proceedings (Mental Health).”

- Data collected from this category should not be considered when assessing the quality of the forensic evaluation.
- G: Guidelines
 - This question is gathering information directly tied to a practice or guideline as indicated by the “Guidelines for Expert Evaluations in Child Abuse/Neglect Proceedings (Mental Health).”
 - For more information about where each question is located in the guidelines please reference the indicated sub-section of the guidelines
 - General Competencies
 - Procedural Guidelines
 - Best Practices

Block I: Evaluator & Evaluation Detail Sheet

B:1.1 – Date of Review

- Format (MM/DD/YYYY)
- Date the Peer Reviewer completed a review of the current evaluation being reviewed.

B:1.2 – Case Identification Number

- Will be provided for each case by the Forensic Coordination Center.

B:1.3 – Local Office Location

- County of origin for the current evaluation being reviewed.

B:1.4 – Evaluation conducted by a

- Selection Options:
 - Private Provider
 - Regional Diagnostic and Treatment Center (RDTC)

B:1.5 Reviewer

- Name of the Peer Reviewer conducting the review of the current evaluation.

E:1.6 – Is the DCP&P Case Goal stated in the Evaluation or Referral Page?

- Case Goal - (CP&P Case Goal): Defines and guides CP&P in its provision of services to each child and family member in need of services.
- See page 36 of the Glossary for a list of CP&P Case Goals.

Block II: Referral Information

B:2.1 – Referral Information Obtained from SAR

B:2.2 – Subject of Assessment

- Who is the subject of the current forensic evaluation?
 - Child
 - Adult

B:2.3 – Subject of Assessment’s Gender

- Specify the gender of the current participant as indicated by the Evaluator.
 - Male
 - Female
 - Transgender
 - Other
 - Specify in the text box provided
 - Not Specified
- This is the same individual as indicated by Q2.2

B:2.4 – Date of Birth

- Format (MM/DD/YYYY)
- Indicate the Date of Birth of the subject of the assessment.
- This is the same individual as indicated by Q2.2
- If no date listed, select “Not specified”

E:2.5 – Indicate the Referral Question/Statement(s) by selecting all that apply

- Referral Question - Addresses the questions, issues, and concerns that are prompting the referral. A short, specific written question/s or statement/s sent to the evaluator by the requesting agency regarding specific mental health concerns that need to be answered directly related to the case.
 - May be written as a question or statement.
- Selection Options:
 - Assess current level(s) of cognitive functioning
 - Assess current level(s) of psychological/ emotional/ behavioral functioning
 - Assess ability to provide adequate care and protection to child/ assess parenting abilities
 - Service needs/ treatment recommendation(s)
 - Other
 - Specify in the text box provided
 - No referral Question/Statement

G:2.7 – Referral question(s) are clearly stated?

- Selection Options:
 - If “Yes” proceed to question 2.9
 - If “No” proceed to question 2.8
- Procedural Guidelines, Page 8

G:2.8 – If "No" was selected for question 2.7 "Referral questions are clearly stated", please briefly explain your selection.

- Make sure to indicate whether the referral question(s) were either not stated or unclear.
- Procedural Guidelines, Page 8

E:2.9 – What evaluation questions are addressed in the words of the Evaluator?

- Questions made by the Evaluator that he or she is looking to answer during the evaluation process.
- These questions are separate from the DCP&P referral questions and case goal.
- Example:
 - Do important parenting deficits exist and are there indications that John Doe may have suffered significantly from these parental deficits?
 - What is Ms. Doe’s level of empathic understanding of John Doe? How is this demonstrated?

Block III: Case Detail Sheet

E:3.1 – Purpose of Evaluation (as indicated by the Evaluator)

- Investigation - Forensic evaluation during the investigatory phase of the case. Most often include allegations of sexual abuse and emotional abuse/neglect. May assist CP&P in determining the impact of an event on a child’s psychological functioning.
- Permanency Planning / Hearing - Requires at least two visits and includes: a clinical interview of each caregiver and an observation of each caregiver with children involved in the case. Should contain recommendations regarding placement and services.
 - If selected answer Q3.1 Type of Permanency Planning / Hearing Evaluation:
 - Interim - Meant to guide reasonable efforts for reunification.
 - Ten Month - Used to prepare a permanency plan for the child or youth in out-of-home placement.
 - Periodic - Evaluation of imminent concerns is used to assess any risks or challenges that the child may incur during the course of protective services or guardianship litigation.
 - Not Specified.
- Litigation of Guardianship - Forensic evaluation during trial preparation after a guardianship complaint has been filed. Consists of fitness and bonding assessments.
- Other – Specify.
- Not Specified / Unclear – If the Evaluator has not indicated the purpose of Evaluation.

B:3.2 – Where is the child(ren) currently placed at the time of this evaluation?

- May select multiple choices if applicable.
- Selection Options:
 - Biological Parent(s)
 - Adoptive / Resource Parent(s)
 - Other – specify

B:3.3 – Type of Assessment for Child (as indicated by the Evaluator).

- If Q2.1 Subject of Assessment “Child” is selected.
- May select multiple choices if applicable.
- Types of Assessment
 - Abuse or Neglect
 - Fire Setting
 - Psychiatric
 - Bonding
 - Psychological

- Sexual Abuse
- Substance Abuse
- Other – Specify

B3.4 – Type of Assessment for Adult (as indicated by the Evaluator)

- If Q2.1 Subject of Assessment “Adult” is selected.
- May select multiple choices if applicable.
- Types of Assessment
 - Psychological
 - Psychiatric
 - Parental Evaluation/Parenting Capacity
 - Bonding
 - Substance Abuse
 - Domestic Violence
 - Sexual Abuse
 - Other – Specify

B3.5 – Qualification of Evaluator

- May select multiple choices if applicable.
- Qualifications
 - Licensed Psychologist – Pd.D. / Psy.D. / Ed.D
 - Psychiatrist
 - MD / DO
 - Licensed LPC
 - LCSW
 - Graduate Level Intern
 - Other – specify

Block IV: Background & Demographic Information

Q4.1 Interview Observations:

Instructions: Next to each item select: Yes - if the practice is described as part of the interview or No - if the practice is not described as part of the interview.

B:4.2 – Subject of Assessment’s Race (as indicated by the Evaluator).

- Asian or Asian Indian
- Black or African American
- Pacific Islander
- White
- Two or More Races
- Other Race – specify
- Not Specified

B:4.3 – Is the Subject of Hispanic or Latino origin (as indicated by the Evaluator).

- For this Tool, Hispanic or Latino origins are not races, please answer separately from Q4.2

B4.4 – Does the Evaluator specify the ethnicity of the subject?

- Answer separately from Q4.2 & Q4.3
- For the purpose of this tool ethnicity is defined as – the fact or state of belonging to a social group that has a common national or cultural tradition
- Example:
 - The subject of the assessment may indicate to the Evaluator that they are of
 - “Black or African American race, non-Hispanic origin, and are Haitian”.

B:4.5 – Does the Evaluator specify the primary language of the subject?

- If “Yes” proceed to question 4.6
- If “No” proceed to question 4.9

B:4.6 – What language was specified?

- English
- Spanish
- Chinese
- Polish
- Korean
- French
- Arabic

- Russian
- Hindi
- Other – specify

G:4.7 – Does the Evaluator indicate how the evaluation was conducted?

- Selection Options:
 - Through an interpreter
 - By the Evaluator – conducted using subject’s primary language
 - Not Specified
- General Competencies, Page 6

G:4.8 – Does the Evaluator avoid using biased language?

- Biased Language – Recognizing that differences should be mentioned only when relevant. Including but not limited to gender, marital status, sexual orientation, racial and ethnic identity, or the fact that a person has a disability should not be mentioned gratuitously¹.
- Example:
 - On the Wechsler series of intelligence tests, the difference in mean scores for Black and White Americans hovers around 15 points. If this figure represents a true difference between the two groups, the tests are not biased. If, however, the difference is due to systematic underestimation of the intelligence of Black Americans or overestimation of the intelligence of White Americans, the tests are said to be culturally biased.²
- General Competencies, Page 7

G:4.9 – Does the Evaluator avoid multiple relationships / conflicts of interests?

- Multiple Relationships / Conflicts of Interest - Multiple relationships occur when a psychologist or licensed mental health professional is in a professional role with a client and:
 - (1) At the same time is in another role (e.g. personal, professional) with the same client.
 - (2) At the same time is in a relationship with a person closely associated with or related to the client.
 - (3) Promises to enter into another relationship in the future with the client or persons associated with said client. Multiple relationships can present a conflict of interest by impairing the objectivity, competence, or effectiveness of the clinician. ³

¹ NJDCF, 2012

² Reynolds & Suzuki

³ NJDCF, 2012

- Example:
 - A psychologist who conducts a forensic evaluation and provides individual therapy for a client.
- General Competencies, Page 7

G:4.10 – Does the summary contain direct quotes that describe the perceptions of the person being evaluated?

- Best Practices, Page 11

Block V: Interview with a Child

B:5.1 – Was an interview with a child conducted?

- A clinical interview where the child is present for the current interview and the Evaluator is not referencing information obtained about the child from a previously conducted interview, observation, or evaluation.
- Responses to questions located in Block V: Interview with a Child should only be obtained from the corresponding section of the forensic evaluation, peer reviewers should not use information from other sections to answer the block questions.
- If “Yes” proceed to question 5.3
- If “No” proceed to question 6.1

Q5.2 Interview with a Child

Instructions: Next to each item select: Yes - if the practice is described as part of the interview or No - if the practice is not described as part of the interview.

G:5.3 – Does the Evaluator describe establishing any of the following with the child?

- Select all that apply
- Selection Options:
 - Ground Rules of the Evaluation
 - Nature and scope of the Evaluation
 - Competency
- Ground rules – rules regarding what a client should expect throughout the evaluation they are participating in, as well as what is expected of the client throughout the particular evaluation.
 - Evaluators inform clients how the evaluation will be used and who will have access to the finalized evaluation report⁴.
- Nature of the evaluation – the purpose of the evaluation and specific questions to be addressed in the evaluation.
- Scope of the evaluation – is determined by the referral or by the court, in consultation with the Evaluator⁵.
- Competency – Whether the child knows the difference between the truth and a lie, real or pretend⁶.
- Best Practices, Page 11

G:5.4 – Does the Evaluator include the child’s version of the situation?

⁴ NJDCF, 2012

⁵ NJDCF, 2012

⁶ NJDCF, 2012

- Using the child's words – may include direct quotes made by the child to describe the incident.
- Best Practices, Page 11

G:5.5 – Does the Evaluator describe the child's affect during the interview?

- Affect – an expressed or observed emotional response
- Example:
 - This Evaluator noted that the child had a constricted affect.
- Best Practices, Page 11

G:5.6 – Does the Evaluator gather a history?

- History – a prior record or incidence with either DCP&P, the police, or the courts.
- Question 5.6 does not refer to the evaluator collecting background information on the child or the family.
- Best Practices, Page 11

G:5.7 – Does the Evaluator note and/or assess for any disabilities of the child?

- Disability - Physical, cognitive, intellectual, or mental condition that:
 - (1) Causes significant impairment in social, occupational, or other important areas of functioning.
 - (2) Limits facets of daily living.
 - (3) Limits an individual's physical functioning, mobility, or dexterity.
- Selection Options:
 - Yes-disabilities noted – a disability was indicated by the Evaluator and appropriate accommodations were indicated as being used by the Evaluator.
 - Yes-no disabilities noted
 - No- no disabilities assessed or noted by the Evaluator
- General Competencies, Page 6 & Best Practices, Page 11

E:5.8 – Does the Evaluator use age appropriate terms?

- Age appropriate terms – terms and language suitable for the child's specific age and his developmental milestones including the child's cognitive and language/communication skills.

G:5.9 – Does the Evaluator have the child describe any of the following relationships?

- Check all that apply
- Selection options:
 - Family Relationships
 - Peer Relationships

- School Relationships
- Other
 - Specify in the text box provided
- Best Practices, Page 11

E:5.10 – Does the Evaluator describe any other domains of the child?

- If “Yes” specify in Reviewer Section Comments

Block VI: Interview with an Adult

B:6.1 – Was an interview with an Adult conducted?

- A clinical interview where the Adult is present for the current interview and the Evaluator is not referencing information obtained about the Adult from a previously conducted interview, observation, or evaluation.
- Responses to questions located in Block VI: Interview with an Adult should only be obtained from the corresponding section of the forensic evaluation, peer reviewers should not use information from other sections to answer the block questions.
- If “Yes” proceed to question 6.2
- If “No” proceed to question 7.1

B:6.2 – Is the Adult a(n):

- Biological Parent
- Adoptive / Resource Parent
- Other – specify

B:6.3 – Is the Adult the subject of the interview or is the interview a collateral interview?

- Subject of the Interview
- Collateral Interview
 - A collateral interview is with a person who is not the subject of the interview.
 - If selected, only respond to question 6.5

Q:6.4 Interviews with an Adult

Instructions: Next to each item select: Yes – if the practice is described as part of the interview or No – if the practice is not described as part of the interview

G:6.5 – Does the Evaluator describe establishing any of the following with the Adult?

- Select all that apply
- Selection Options:
 - Ground Rules
 - Nature and scope of the Evaluation
 - Informed Consent
- Ground rules – rules regarding what a client should expect throughout the evaluation they are participating in, as well as what is expected of the client throughout the particular evaluation.
 - Evaluators inform clients how the evaluation will be used and who will have access to the finalized evaluation report⁷.

⁷ NJDCF, 2012

- Nature of the evaluation – the purpose of the evaluation and specific questions to be addressed in the evaluation.
- Scope of the evaluation – is determined by the referral or by the court, in consultation with the Evaluator⁸.
- Informed Consent - the purpose of the research, expected duration, and procedures.
- Procedural Guidelines, Page 8 & Best Practices, Page 11

G:6.6 – Does the Evaluator describe the Adult’s affect during the interview?

- Affect – an expressed or observed emotional response.
- Example:
 - This Evaluator noted that the Adult had a flat affect.
- Best Practices, Page 11

G:6.7 – Does the Evaluator note and/or assess for any disabilities of the Adult?

- Disability - Physical, cognitive, intellectual, or mental condition that:
 - (1) Causes significant impairment in social, occupational, or other important areas of functioning.
 - (2) Limits facets of daily living.
 - (3) Limits an individual’s physical functioning, mobility, or dexterity.
- Selection Options:
 - Yes-disabilities noted – a disability was indicated by the Evaluator and appropriate accommodations were indicated as being used by the Evaluator.
 - Yes-no disabilities noted
 - No- no disabilities assessed or noted by the Evaluator
- General Competencies, Page 6 & Best Practices, Page 11 and 15

G:6.8 – Does the Evaluator describe any of the following histories of the Adult?

- Check all that apply
- Selection Options:
 - Education History
 - Work History
 - Relationship History
 - Family History
 - Other History
 - Specify
- Best Practices, Page 11

G:6.9 – Does the Evaluator describe the family relationships of the Adult?

⁸ NJDCF, 2012

- Example:
 - Ms. Doe reported her current relationship with her family; she said her mother was supportive and would help with the children.
- Best Practices, Page 15

G:6.10 – Does the Evaluator describe the parenting style of the Adult?

- Parenting style – the emotional climate in which parents raise their children⁹.
- Example:
 - Ms. Doe acknowledged that she is not 100% involved. She reported that she loves her children but does not show it.
- General Competencies, Page 5

G:6.11 – Does the Evaluator describe the parenting capacity of the Adult?

- Example:
 - When asked about current discipline, she felt she could verbally control Jane Doe.
- General Competencies, Page 5

E:6.12 – Does the Evaluator explore personal skills of the Adult?

- This question pertains to the current allegation(s) only.
- Example:
 - Ms. Doe was a calm and rational individual. She was defensive, but within the limits of most individuals in this setting.

G:6.13 – Does the Evaluator describe risk factors of the Adult?

- Risk factors – based on empirical evidence that they are related to increased risk of abuse or neglect¹⁰.
- Example:
 - Pattern of domestic violence.
 - Grossly impaired judgment and insight.
- Best Practices, Page 12

G:6.14 – Does the Evaluator describe functional abilities of the Adult?

- This question pertains to the current allegation(s) only.
- Functional Abilities - Ability to perform daily living skills such as feeding, bathing, dressing and other independent living skills such as grocery shopping and cleaning the house.

⁹ Darling and Steinberg, 1993

¹⁰ NJ Child Welfare Training Partnership, 2015

- Example:
 - She reported owning her own home and she and her fiancée were working two jobs and financially stable.
- General Competencies, Page 6

G:6.15 – Does the Evaluator take into consideration other relatives if or when a parent is unable to meet the needs of the child?

- This question pertains to the current allegation(s) only.
- Example:
 - It appeared that a relative placement was not successful as one of Jane Doe's siblings had a previous legal history. As a result, Jane Doe was placed in a shelter.
- General Competencies, Page 6

Block VII: Psychological Inventories and Interpretation

B:7.1 – Were tools (psychometric measures / scales / inventories) used to complete the evaluation?

- No tools were used; interview only.
 - If selected proceed to question 8.1
- Tools were used.
 - If selected proceed to question 7.2

B:7.2 – Indicate which (if any) Cognitive and Achievement Test(s) were used.

- Adaptive Behavior Assessment System (ABAS)
- General Ability Measure for Adults (GAMA)
- Kaufman Brief Intelligence Test (KBIT)
- Shipley Institute of Living Scale
- Stanford-Binet
- Test of Nonverbal Intelligence (Toni)
- Vineland Adaptive Behavior Scales
- Wechsler Abbreviated Scale of Intelligence (WASI)
- Wechsler Adult Intelligence Scale (WAIS)
- Wechsler Individual Achievement Test (WIAT)
- Wechsler Intelligence Scale for Children (WISC)
- Wechsler Memory Scale (WMS)
- Wechsler Preschool and Primary Scale of Intelligence (WPPSI)
- Wide Range Achievement Test (WRAT)
- Woodcock Johnson
- Universal Nonverbal Intelligence Test
- Other – specify
- None

B:7.3 – Indicate which (if any) Multiscale Inventories were used.

- Behavior Assessment System for Children (BASC)
- Child Behavior Checklist
- Conners Behavior Rating
- Millon Adolescent Personality Inventory (MAPI)
- Millon Clinical Multiaxial Inventory (MCMI)
- Minnesota Multiphasic Personality Inventory (MMPI)
- Personality Assessment Inventory (PAI)
- Personality Assessment Screener (PAS)
- Sixteen Personality Factors Questionnaire

- Youth Self Report
- Other – specify
- None

B:7.4 – Indicate which (if any) Clinical and/or Personality Scale(s) were.

- Battelle Developmental Inventory
- Beck Depression Inventory
- Beck Anxiety Inventory
- Beck Hopelessness Scale
- Beck Youth Inventory
- Children’s Depression Inventory
- Conners ADHD
- Projective Drawings
- Psychiatric Diagnostic Question
- Rorschach
- Sentence Completion
- Other – specify
- None

B:7.5 – Indicate which (if any) Psychological Tests in Child-Related Forensic Issues were used.

- Adult Adolescent Parenting Inventory
- Child Abuse Potential Inventory
- Child Behavior Checklist
- Child Sexual Behavior Inventory
- Millon Adolescent Clinical Inventory
- Parent-Child Relationship Inventory
- Parenting Stress Index
- Trauma Systems Checklist
- Other – specify
- None

B:7.6 – Indicate if any Neurological Tools were Used

B:7.7 – Indicate and describe other tools that were used to complete the evaluation (Excluding Neurological Tools)

G:7.8 – Does the Evaluator adequately describe the purpose of the psychometric tools used in the Evaluation?

- Example:

- The Conners BRS is a multidimensional rating scale commonly used to evaluate a broad range of psychological concerns.
- General Competencies, page 7

G:7.9 – Does the Evaluator describe how to interpret the results of the psychometric tools used in the Evaluation?

- Example
 - The Conners BRS: A “Very Elevated” score indicated many more concerns than are typically reported.
- General Competencies, page 7

G:7.10 – Does the Evaluator use the current version(s) of the psychological tools used in the Evaluation?

G:7.11 – Does the Evaluator describe any threats to the validity of the results?

- Selection Options:
 - Yes – A threat(s) to the validity of the psychometric tools were indicated/described by the Evaluator.
 - No – No threat(s) to the validity of the psychometric tools were indicated/described by the Evaluator.
 - No, but threat(s) to validity are present – The Evaluator did not indicate/describe any threats to the validity of the psychometric tools, however, a threat does exist.
- Validity of Psychological Tools – refers to how well the assessment tool measures the underlying outcome of interest¹¹.
- Threats to the validity of a psychometric test may include the following:
 - Not adequately reporting the scores
 - Use of a translator
 - Conducting the test in a language other than the individual’s native language
- Example:
 - Ms. Doe completed the Parent Version of the Conners BRS. Results of the validity scales indicated that Ms. Doe responded to the items consistently, but also indicated a negative response style. Results should be interpreted with caution.
- General Competencies, Page 6

E:7.12 - Indicate the threat(s) to the validity of the results (Whether they were reported by the Evaluator or not).

¹¹ Sullivan, 2011

- Selection Options
 - Issues pertaining to cultural factors
 - May include issues related to language or population norms of the tool used
 - Issues pertaining to participant characteristics
 - May include issues related to physical factors, such as a disability, or intellectual ability
 - Issues pertaining to test administration
 - May include issues related to non-standardized test use or inappropriate for the reason for referral
 - Issues pertaining to validity indexes
 - May include issues related to faking good/ overly defensive subject or overly negative/ malingering.
 - Other threats to validity

Block VIII: Evaluation Questions

Q8.1 - Instructions: Next to each item select: Yes - if the practice is described as part of the interview or No - if the practice is not described as part of the interview.

G:8.2 – Does the Evaluator assess the risks that need to be addressed, as indicated by the referral?

- Risks – matters where a child’s health and/or welfare may have been harmed or have the potential to be harmed.¹²
- Example:
 - This evaluation does find evidence of significant psychopathology that is expected to interfere with safe and effective parenting at this time.
- If questions fall outside the scope of the current evaluation, as indicated by the referral, select “N/A.”
- Best Practices, Page 12

G:8.3 – Does the Evaluator describe what progress has been made towards eliminating the risks/harm, as indicated by the referral?

- Example:
 - At this time this Evaluator notes no observable improvement in her functioning or ability to protect the children and provide for their basic needs.
- If questions falls outside the scope of the current evaluation, as indicated by the referral, select “N/A.”
- Best Practices, Page 12

G8.4 – Does the Evaluator identify the impact of the presenting problem, as indicated by the referral question?

- Impact – identification of how the presenting problem has affected the particular psychological, behavioral, and developmental needs of the child.¹³
- Example:
 - Despite the noted parenting deficits, it is unlikely that the child suffered significant harm; it is more likely that the parent-child relationship was strained as a result.
- Best Practices, Page 13

¹² NJDCF, 2012

¹³ NJDCF, 2012

G:8.5 – Does the Evaluator, within the scope of their professional judgment, identify the impact of the child’s history of abuse / neglect?

- Professional judgment – contains relevant professionally sound observations, results, and opinions in matters where a child’s health or welfare may have been harmed or placed at risk.¹⁴
- Example:
 - Her pattern of erratic and dysfunctional behavior continues, including her recent arrest for assault and her decision not to go home and take care of the children and continue her pathological relationship with Mr. Doe.
- Best Practices, Page 12

E:8.6 – Does the Evaluator identify additional steps needed to address the risks / harm?

- Example:
 - A complete review of Ms. Doe’s psychological report would likely support or oppose the findings of this Evaluator.

G:8.7 – Does the Evaluator identify areas of strength?

- Example:
 - Stable Housing.
 - Currently complying with medications.
- Guidelines, Page 5 & Best Practices, Page 12

G:8.8 – Does the Evaluator identify new areas of need not previously identified?

- New areas of need may include - risks or harm which may not have been identified in the initial investigation or the referral provided by DCP&P.
- If “Yes” briefly describe the new areas in the text box provided.
- Best Practices, Page 12

E:8.9 – Does the Evaluator list and / or describe services that are currently provided?

- If no services are currently provided select “N/A”.

¹⁴ NJDCF, 2012

Block IX: Collateral Information

Q9.1 **Collateral Information** - Was the following information accessed by the Evaluator or were attempts made to access?

Instructions: Please check off next to each item if the Interview, Record, or History was accessed by the Evaluator.

E:9.2 – Which of the following Collateral Interviews were accessed by the Evaluator?

- Collateral Interview - contacts with persons or third parties who may provide relevant information to address the evaluation's referral questions. A collateral interview is with a person who is not the subject of the evaluation. These may include interviews with teachers, co-workers, doctors, therapists, resource parents, family members, etc.
- May Include:
 - Interview with Child
 - Interview with Parent(s)
 - Interview with Spouse/Partner
 - Interview with other Relative(s)
 - Interview with Resource Family
 - Interview with Teacher(s) or other school personnel
 - Taped Interviews
 - Interview with other collateral – specify

E:9.3 – Which of the following Collateral Records and/or Histories were accessed by the Evaluator?

- Collateral Records and/or Histories - Documentation including reports from schools, health care providers, previous evaluations, and relevant legal documents that provide pertinent information to address the evaluation's referral question(s). An evaluator may review collateral documents to gather important historical data and substantiate psychological, psychiatric, or behavioral issues noted elsewhere.
- May Include:
 - Medical Records
 - CP&P Records
 - School Records
 - Prior Evaluations and/or Observations
 - Law Enforcement Records
 - History of Offenses
 - Court Complaint
 - Mental Health / Psychiatric History
 - Substance Abuse History

- Domestic Violence History
- Prior Complaints Filed
- Other Records – specify

G:9.4 – To what extent were appropriate background materials provided by DCP&P to address the purpose of the referral?

- No background material provided
- Insufficient
- Sufficient
- Procedural Guidelines, Page 8 & Best Practices, Page 13

G:9.5 – Please explain your previous selection

- Pertains to question 9.4

Block X: Diagnosis

E: 10.1 – Does the Evaluator conduct a diagnostic assessment on the Subject of the Evaluation?

- Diagnostic Assessment – can include exploration, discussion or testing of an individual for a possible diagnosis.
- If yes, proceed to question 10.2
- If no, proceed to question 10.4

E: 10.2 – Does the Evaluator provide a DSM diagnosis (or ICD10) for the Subject of the Evaluation?

- Selection Options:
 - Yes, by the Evaluator – A formal diagnosis of the individual was provided by the Evaluator based on the current assessment
 - Diagnosis - Based on the signs, symptoms, and evaluative findings of a psychological assessment or evaluation, e.g., DSM IV/V or ICD-9/10 codes.
 - If the Evaluator indicated that the person does not meet criteria for a diagnosis select “Yes, by Evaluator”
 - Yes, by History - A formal diagnosis of the individual was provided but it was indicated as “By History.”
 - No – A formal diagnosis of the individual was not provided.

E: 10.3 – Was there reasonably sufficient evidence or basis for the diagnosis made?

E: 10.4 – Should the Evaluator have conducted a diagnostic assessment?

- Yes
- No
- N/A for the evaluation
- If Yes or N/A is selected please explain why.

Block XI: Summary

B:11.1 – Was a summary/ conclusions/ or diagnostic impression included?

- If “Yes” proceed to question 11.3
- If “No” proceed to question 12.1

Q11.2 Summary of Findings

Instructions: Next to each item select: Yes - if the practice is described as part of the interview or No - if the practice is not described as part of the interview

G:11.3 – Does the Evaluator provide a background summary?

- Best Practices, Page 11 & 16

G:11.4 – Does the summary address referral questions?

- Best Practices, Page 9

G:11.5 – Does the Evaluator describe the nature of the allegations?

- Best Practices, Page 11

G:11.6 – If applicable, does the Evaluator indicate clinical findings in the summary?

- Clinical Findings – interpretive statements based upon a client’s medical and/or psychological symptomatology as well as previous and current evaluative data.
- Best Practices, Page 11

G:11.7 – Does the Evaluator identify clinical interventions?

- Clinical Interventions - Counseling or psychotherapy services provided to clients to address psychological, emotional, or behavioral problems as well as stressful life events that might impair his/ her ability to function or thrive. Emphasis is on implementing services which increase client’s capacity to adaptively cope, manage, or modify symptoms, behaviors, and maladaptive coping mechanisms.
- General Principles and Guidelines, Page 5

Block XII: Recommendations

B:12.1 – Were recommendations made?

- If “Yes” proceed to question 12.3
- If “No” proceed to question 13.1

Q12.2 Recommendations

Instructions: Next to each item select: Yes - if the practice is described as part of the interview or No - if the practice is not described as part of the interview.

G:12.3 – Does the Evaluator make recommendations to address the needs of the child described in the referral, regardless of the subject of the evaluation?

- Example:
 - At this point, the children’s right to have some chance of permanency and a stable, nurturing environment outweighs Ms. Doe’s right to have another round of services. I strongly recommend against reunification and for termination of parental rights.
- General Competencies Guidelines, Page 5

G:12.4 – Does the Evaluator make recommendations that were relevant to the purpose of DCP&P?

- General Competencies, Page 5

G:12.5 – Does the Evaluator make recommendations to address the risk and / or harm that include:

- Selection Options:
 - Services that are evidence based
 - Services that are generally accepted in clinical practice as appropriate for use
 - Services that are not evidence based, but are evidence informed or promising programs
 - Services for which there is either weak or no evidence
 - It is uncertain whether some of the recommendations are supported by current scientific evidence
- May select multiple choices if applicable.
- Evidence based - Best practice, based on a thorough evaluation of evidence from published research studies that identify interventions to maximize the chance of benefit, minimize the risk of harm, and deliver treatment at an acceptable cost. Evaluators should utilize research literature and experts’ findings when conducting

interviews, testing measures, interpreting results, and when making recommendations for treatment or services.

- General Competencies, Page 6

E:12.6 – Does the Evaluator make recommendations that include a description of expected outcomes from interventions and/or treatments?

- Example:
 - The child should be referred for Trauma-Focused-CBT (TF-CBT) to address their exposure to domestic violence and physical abuse. TF-CBT is targeted to treat trauma-induced symptoms as well as assist the child in developing prosocial and effective coping skills.

E:12.7 – Does the Evaluator make recommendations that describe conditions most likely to yield successful outcomes?

- Example:
 - Ms. Doe should be referred to a psychiatrist who has experience treating Bipolar Disorder. Should Ms. Doe comply with this service and demonstrate a strong compliance with her medication (over several months) as well as improved insight and coping skills, she could begin to have some unsupervised role with Jane Doe.

E:12.8 – Does the Evaluator include a hierarchy timeline for service based recommendations to be completed?

- If “Yes” proceed to question 12.9
- If “No” proceed to question 12.10

E:12.9 – If the Evaluator included a hierarchy timeline for service based recommendations do they indicate which services are to occur concurrently?

E:12.10 – Does the Evaluator make recommendations that describe the degree to which specific interventions/treatment are likely to be successful?

- Example:
 - I cannot identify any service that has a realistic chance of changing the risk factors in this case within a reasonable time frame to protect the children and provide permanency.

E:12.11 – Does the Evaluator make recommendations that take into account the individual’s race, ethnicity, gender or other self-defining characteristics?

E12.12 – How many clinical or program based service recommendations were specified by the Evaluator?

- Selection Options¹⁵
 - 0
 - 1 – 2
 - 3 – 5
 - > 5

E12.13 – How many family focused recommendations were specified by the Evaluator?

- Example:
 - It is recommended that Jane Doe be referred to a mentoring program for a Big Sister as a source of support and to assist Jane in strengthening her sense of self.
- Selection Options
 - 0
 - 1 – 2
 - 3 – 5
 - > 5

¹⁵ NJ Child Welfare Training Partnership, 2015

Block XIII: Evaluation Rubric

For the following questions please indicate the degree to which the evaluation/evaluator meets the following criteria, using the scale

- 1 – Strongly Disagree
- 2 – Disagree
- 3 – Somewhat Disagree
- 4 – Somewhat agree
- 5 – Agree
- 6 – Strongly Agree

R:13.11 – To what degree does the evaluation contain relevant, professional sound observations, results, and opinions

R:13.12 – To what degree does the evaluation address the particular psychological, behavioral, and developmental needs of the child and/or parent(s)

R:13.13 – Does the Evaluator include any of the following (Check all that apply):

- Does the Evaluator describe the degree to which severe and enduring harm would occur if the child is removed from their current placement?
- Does the Evaluator describe the degree to which the parent(s) are fit and able to parent the child?
 - Biological, Adoptive, or Resource parent
- Does the Evaluator describe the degree to which the parent(s) can mitigate harm?
 - Biological, Adoptive, or Resource parent
- Does the Evaluator describe the degree of the relationship between the child and the parent(s)?
 - Biological, Adoptive, or Resource parent
- Does the Evaluator describe or make recommendations to change the permanency plan?
- Does the Evaluator describe an appropriate visitation plan between the parent(s) and the child?
- Does the Evaluator describe the services needed for reunification?

R:13.14 – To what degree does the Evaluator use a multimodal approach to draw conclusions about the current case

- Multimodal - Use of multiple sources. The Evaluator does not rely on or make conclusions based on one source, e.g., records, observations, psychological testing.

R:13.15 – To what degree does the Evaluator take into account the cultural norms of the child and/or parent(s) being evaluated

- Cultural competence - A set of attitudes, behaviors, and policies that integrates knowledge about groups of people into practices and standards to enhance the quality of services to all groups served.
 - Also, the ability of individuals and systems to respond respectfully and effectively to people of all cultures, classes, races, ethnic backgrounds, sexual orientations, and faiths or religions in a manner that recognizes, affirms and values the worth of individuals, families, tribes, and communities, and protects and preserves the dignity of each.

R:13.16 - Does the Evaluator explore any of the following (Check all that apply):

- Explore cultural explanations of the allegations?
- Explore the client's cultural background and ways it might influence their behavior?
- Discuss the client's cultural background and ways it might influence their parenting capacity?

R:13.17 – To what degree does the Evaluator display expertise with the instruments employed, including psychological and intelligence tests

R:13.18 – For a bonding evaluation – To what degree does the Evaluator attempt to address the best placement for the needs of the child and/or address consequences of removing the child from resource parent(s) vs biological parent(s)

- Only appropriate if either question 3.4 “Bonding” assessment was selected or if question 10.1 “Yes” was selected

R:13.19 – To what degree does the summary outline and address the problems stated in the referral section

R:13.20 – To what degree do the recommendations promote the psychological and physical well-being of the child, and when appropriate, facilitate the safe reunification of the child with the parent

R:13.21 – To what degree are the recommendations made by the Evaluator tied to observable outcomes

R:13.22 – Indicate the overall quality of the forensic evaluation written by the Evaluator

R:13.23 – Is a secondary review of this case needed

- May be requested if issues of reliability or clarity are raised and a secondary opinion may help

Glossary of Terms

Biased Language: Recognizing that differences should be mentioned only when relevant. Marital status, sexual orientation, racial and ethnic identity, or the fact that a person has a disability should not be mentioned gratuitously.

CASA: Court Appointed Special Advocates are specially trained community volunteers appointed by a judge to advocate on behalf of children in out-of-home placements to ensure their well-being and ultimate placement in safe and nurturing permanent homes.

Case Detail Sheet: Assist the Office of Quality in preparing the final report for the county being reviewed. These reports provide the necessary information to assist the county in developing their Program Improvement Plan. They include a summary of the family picture and a rationale for each score.

Case Goal (CP&P Case Goal): Defines and guides CP &P in its provision of services to each child and family member in need of services and may include one of the following:

Maintenance in Own Home/Family Stabilization: Keeping a child in his or her home when the circumstances do not necessitate removal from the home or keeping the child in the home to which he or she was returned or placed following out-of-home placement, regardless of the child's biological or legal ties to the person or persons, when reunification with the parent(s) has been ruled out.

Reunification (Return Home): This is the case goal when the child is in any type of substitute care. Case activities are directed toward safely returning the child to, or placing the child with, a parent when the circumstances necessitating out-of-home placement have been resolved and the parent has expressed an interest in, and displayed the willingness and ability to, care for the child with support services if necessary.

Adoption: Adoption is the legal transfer of all parental rights and responsibilities from the birth and/or legal parent to another person who desires to assume those rights and responsibilities. The goal of Adoption is the first and best choice for a child who cannot return home because it provides children with the highest level of legal and emotional security. Types of Adoption include: Relative, Family Friend, Foster Home, or Selected Home.

Kinship Legal Guardian: A caregiver who is willing to assume care of a child due to parental incapacity, with the intent to raise the child to adulthood, and who is appointed the kinship legal guardian of the child by the court. This person shall be responsible for the care and protection of the child and for providing for the child's health, education, and maintenance. Kinship Legal Guardianship (KLG) is intended to be used when "adoption of the child is neither feasible nor likely." KLG is intended to be permanent and self-sustaining, as evidenced by the transfer of certain parental rights to the caregiver, while the parent retains the right to consent to adoption, an obligation to pay child support, and the right to have ongoing contact with the child.

Independent Living: Only appropriate for adolescents 16 to 18 years of age, when there is absolutely no alternative. Case activities are directed toward the achievement and maintenance of an adolescent in a living arrangement that allows him or her to eventually

function on his or her own.

Other Long-Term Specialized Care: In very rare cases, the case activities are directed toward the placement or maintenance of a child in a long term, specialized care living arrangement. The case goal is chosen when no appropriate family is able or willing to care for, and meet the needs of, a child with a serious medical, physical, emotional, or mental disability, and the child will remain institutionalized because no less restrictive living arrangement can meet his or her needs for care and treatment.

Individual Stabilization: Used only for parent(s) whose children have all been placed out-of-home, and the permanent plan for every child of that parent is other than Reunification. Also used when older adolescents (i.e., young adults), age 18 to 21, who are in, or will soon be transitioned into, an independent living program or setting, agree to continue to receive services from CP&P, and for whom no other goal is appropriate.

Child Abuse or Neglect (FN Docket Type): Involves complaints filed with the court alleging child abuse or neglect. After a CP&P caseworker conducts a preliminary investigation into an allegation of child abuse or neglect, the agency may file a complaint with the Deputy Attorney General (DAG). The DAG then files a Complaint and Order to Show Cause with the Court to protect the child from harm. The Court assigns a return date on the Order to Show Cause within 21 days. A Fact Finding Hearing is to be heard within 4 months. The family may stipulate to a finding of abuse. If a stipulation is not signed, a Fact Finding Trial is held to determine whether or not the child has been abused or neglected.

Child in Placement (FC Docket Type): Established for any child who is placed outside his or her home under the supervision of the CP&P. A caseworker files a Notice of Placement with Children in Court Services within 72 hours of the placement. These cases are reviewed by a trained citizen volunteer board known as the Child Placement Review (CPR) Board who reviews the case to make recommendations to the court for a permanent case plan for the child as soon as possible.

Child Protective Services: The name of governmental agencies responsible for providing child protection as well as monitoring the welfare and safety of children, which includes responding to reports of child abuse or neglect. In New Jersey, the child protection and welfare agency is Child Protection and Permanency (CP&P). CP&P assesses reports of child maltreatment, and if it is determined that the child is at risk or has been abused or neglected, then the agency works to ensure services and supports are offered to the child and family.

Child Protective Services Investigation: Activity of gathering information necessary to make a formal agency determination, to stand up in legal proceeding if necessary, as to whether child abuse or neglect has occurred. Child Protective Services findings include:

Substantiated: The investigation revealed a preponderance of evidence that a child has been abused or neglected because the child was harmed or placed at substantial risk of serious harm by a parent or caregiver. Substantiated findings are disclosed upon a CARI request.

Established: A preponderance of the evidence that a child is an abused or neglected child as defined by definition, but the act or acts committed or omitted do not warrant a finding of substantiation upon consideration of aggravating and mitigating factors. Established findings are not disclosed upon a CARI request but are maintained in agency records.

Not Established: There is not a preponderance of the evidence that the child is an abused or neglected child by definition, but evidence indicates that the child was harmed or placed at risk of harm. Not Established findings are not disclosed upon a CARI request but are maintained in agency records.

Unfounded: There is not a preponderance of evidence that a child was harmed or placed at risk of harm by a parent or caregiver. Unfounded findings are not disclosed upon a CARI request and are eligible for expunction from agency records if no further allegations are made in the next three years.

Clinical Findings/Impressions: Interpretive statements based upon a client's medical and/ or psychological symptomatology as well as previous and current evaluative data.

Clinical Intervention: Counseling or psychotherapy services provided to clients to address psychological, emotional, or behavioral problems as well as stressful life events that might impair his/ her ability to function or thrive. Emphasis is on implementing services which increase client's capacity to adaptively cope, manage, or modify symptoms, behaviors, and maladaptive coping mechanisms.

Collateral Documents: Documentation including reports from schools, health care providers, previous evaluations, and relevant legal documents that provide pertinent information to address the evaluation's referral question(s). An evaluator may review collateral documents to gather important historical data and substantiate psychological, psychiatric, or behavioral issues noted elsewhere.

Collateral Interviews: Contacts with persons or third parties who may provide relevant information to address the evaluation's referral questions. A collateral interview is with a person who is not the subject of the evaluation. These may include interviews with teachers, co-workers, doctors, therapists, resource parents, family members, etc.

Concurrent Planning: A case practice that provides reunification services while simultaneously implementing an alternative or back-up permanency plan in the event that reunification cannot be accomplished within the required time frames. Concurrent, rather than sequential, planning efforts move children more quickly from the uncertainty of resource care to the security of a permanent family.

Cultural Competence: A set of attitudes, behaviors, and policies that integrates knowledge about groups of people into practices and standards to enhance the quality of services to all groups served. Also, the ability of individuals and systems to respond respectfully and effectively to people of all cultures, classes, races, ethnic backgrounds, sexual orientations, and faiths or religions in a manner that recognizes, affirms and values the worth of individuals, families, tribes and communities, and protects and preserves the dignity of each.

Diagnosis: Based on the signs, symptoms, and evaluative findings of a psychological assessment or evaluation, e.g., DSM IV/V or ICD-9/10 codes.

Disabilities/Disability: Physical, cognitive, intellectual, or mental condition that: (1) causes significant impairment in social, occupational, or other important areas of functioning ; (2) limits facets of daily living; (3) or limits an individual's physical functioning, mobility, or dexterity.

Dodd Removal (DODD): Occurs when CP&P determines a child has been harmed or is in ‘imminent risk of harm’ as a result of its investigation. In these circumstances, the law authorizes CP&P to remove a child from the home without a court order; the Division must appear before the judge within two court days.

Evaluations: There are a range of evaluations that may be considered at various stages in child protective services including:

Bonding Evaluation: A specialized type of assessment with the goal to determine the nature of the child’s attachment to birth parents and foster parents, especially to address the question of who occupies the position of greatest centrality in the child’s emotional life. Most commonly used at the end of adoption cases, and is sometimes used without a concurrent psychological evaluation.

Forensic Evaluation: A forensic evaluation in child welfare proceedings and child protective service matters is an evaluation necessary to assist the court and/or CP&P in case planning or to resolve a case. Forensic evaluations are not for the purpose of providing mental health treatment but rather at the request of a court, an attorney, or an administrative body to assist in addressing a forensic referral question. A forensic evaluation may contain recommendations for mental health treatment.

Investigative Evaluation: Forensic evaluation during the investigatory phase of the case. Most often include allegations of sexual abuse and emotional abuse/neglect. May assist CP&P in determining the impact of an event on a child’s psychological functioning.

Litigation for Guardianship Evaluation: Forensic evaluation during trial preparation after a guardianship complaint has been filed. Consists of fitness and bonding assessments.

Parental Capacity Evaluation: Requires at least two visits and includes: a clinical interview of each caregiver and an observation of each caregiver with children involved in the case. Should contain recommendations regarding placement and services.

Permanency Planning/Hearing Evaluation: Forensic evaluation at the time of referral and over the course of a child or adolescent’s time under CP&P. Contributes to the decisions made about placement, permanency, and parent rights.

Interim Evaluation: Meant to guide reasonable efforts for reunification.

Ten Month Conference: Used to prepare a permanency plan for the child or youth in out-of-home placement.

Periodic Evaluation: Evaluation of imminent concerns is used to assess any risks or challenges that the child may incur during the course of protective services or guardianship litigation.

Psychiatric Evaluation: Forensic evaluation conducted by a medical doctor with a MD or DO degree. Reviews functioning of the consumer and can assess if psychotropic medications would assist that individual.

Psychological Evaluation: Forensic evaluation conducted by a PhD, PsyD, or EdD. Can aid in the determination of a diagnosis, identify risk factors, and recommend the best treatments. Also provides information about the person, IQ, mental status, and personality characteristics.

Psychosexual Evaluation: Can be conducted with children and adults.

Forensic evaluation designed to determine:

- The risk of the individual repeating the behavior
- Interventions that will be most effective
- Specific risk factors
- One's willingness to comply with treatment recommendation and interventions
- Identifying factors that may prevent engagement in treatment and interventions
- Identifying strengths and protective factors that are preventative

Psychosexual evaluations do not:

- Determine guilt or innocence
- Identify whether an individual is or is not a "sex offender"
- Conclude whether an adult or juvenile meets the profile of a sex offender

Evidence Based Practice: Best practice, based on a thorough evaluation of evidence from published research studies that identify interventions to maximize the chance of benefit, minimize the risk of harm and deliver treatment at an acceptable cost. Evaluators should utilize research literature and experts' findings when conducting interviews, testing measures, interpreting results, and when making recommendations for treatment or services.

Foster Care: All living arrangements involving a child being placed outside his/her home. May include: homes of relatives, not-related foster parents, or group homes.

Functional Abilities: Ability to perform daily living skills such as feeding, bathing, dressing and other independent living skills such as grocery shopping and cleaning the house.

Ground Rules: Rules regarding what a client should expect throughout the evaluation they are participating in, as well as what is expected of the client throughout the particular evaluation. Evaluators inform clients how the evaluation will be used and who will have access to the finalized evaluation report.

Guardianship and Termination of Parental Rights (FG) Court Case: Guardianship involves the termination of parental rights of the parent or guardian of the child or children. The filing of a Termination of Parental Rights Complaint is often the end result of a proceeding for abuse or neglect.

Inferred Referral Question: When the referral question is unclear or unstated; if the Evaluator specifies that he or she is making an inference about the scope of the evaluation.

Interview with a Child: A clinical interview where the child is the subject of the evaluation.

Interview with an Adult: A clinical interview where the biological parent, foster parent, or other caregiver is the subject of the evaluation.

Law Guardian: Attorney working for the Office of the Public Defender, Office of Law Guardian. Statute mandates that all children in foster care be assigned legal counsel.

Multi-modal: Use of multiple sources. The Evaluator does not rely on or make conclusions based on one source, e.g., records, observations, psychological testing.

Multiple Relationships/Conflicts of Interest: Multiple relationships occur when a psychologist or licensed mental health professional is in a professional role with a client and: (1) at the same time is in another role (e.g. personal, professional) with the same client; (2) at the same time is in a relationship with a person closely associated with or related to the client; or (3) promises to enter into another relationship in the future with the client or persons associated with said client. Multiple relationships can present a conflict of interest by impairing the objectivity, competence, or effectiveness of the clinician. An example of this might be a psychologist who conducts a forensic evaluation and provides individual therapy for a client.

MVR (Minimum Visitation Requirement): The policy and procedures for CP&P Workers to regularly meet with each child, adolescent or young adult, his or her parent, and, if applicable, out-of-home placement provider, for families in open case status, commonly known as Minimum Visitation Requirements (MVR).

Needs Assessment: An assessment conducted to determine and identify the needs or discrepancies between current conditions and desired/ optimal conditions to support the well-being and safety of children and address any deficits that might impair parent/ guardian's ability to adequately care for children.

Referral Question/s: Addresses the questions, issues, and concerns that are prompting the referral. A short, specific written question/s or statement/s sent to the evaluator by the requesting agency regarding specific mental health concerns that need to be answered directly related to the case.

Resource Families/Parents: Relatives, non-relative foster parents, as well as individuals interested in providing care for a child. There are three types of resource family care providers:

Adoptive Caregivers: Provide permanent care for children whose parents have had their parental rights terminated.

Foster Caregivers: Individuals and families who volunteer to become temporary caregivers to children in need of a home due to protective or other social service reasons.

Kinship Caregivers: Related to a child in placement through blood, marriage, civil union, domestic partnership, or adoption. Kinship caregivers may also be connected to the child by an established positive psychological or emotional relationship.

Risk Assessment: An assessment performed to identify the likelihood that a child is at risk for or will be abused or maltreated by a caregiver in the future.

Safety Assessment: An assessment performed to identify any potential threats to a child's wellbeing or safety that might exist in the home.

SAR: Special Approval Request CP&P Form 16-76: The SAR is used to request programmatic services provided by CP&P which require the approval of the Local Office Manager or designee. The form, once signed, serves as documentation to a vendor that provision of service has been approved by CP&P management. The SAR is used to document to a service provider (i.e., a vendor) that payment for the service has been approved by CP&P management.

Title 9: Found at NJSA 9:6-8.21, deals with instances where the child has been abused and/or neglected by a parent or caregiver. Title 9 explicitly defines criminal offenses that deal with abuse, abandonment, cruelty, and neglect of children.

Title 30: Found at NJSA 30:4c-12, deals with instances where a child requires care and supervision by CP&P or other action to ensure the health and safety of the child. Under this law, CP&P may apply to the Family Part of the Chancery Division of the Superior Court for an order making the child a ward of the court and placing the child under the care and supervision or custody of CP&P.

Visitation Supervision Levels: There are four types of supervision for visits between parents and their children when in out-of-home placement:

Therapeutic: Requires a professional to supervise and there is usually a clinical purpose.

Supervised: The parent and child need to be in sight and sound distance from an objective person and no alone time is permitted.

Observed: An objective person is involved but does not have to be in sight and sound distance at all times.

Unsupervised: No supervision is needed.

References

Darling, N. & Steinberg, L. (1993). Parenting style as context: An integrative model.

Psychological Bulletin, 113(3), 487-496.

New Jersey Department of Children and Families. (2012). Guidelines for expert Evaluations in Child Abuse/Neglect Proceedings.

Sullivan, G.M. (2011). A primer on the validity of assessment instruments. *Journal of Graduate Medical Education*, 3(2), 119-120.

Appendix B

New Jersey Department of Children and Families' Guidelines for Expert
Evaluations in Child Abuse/Neglect Proceedings (Mental Health)

New Jersey Department of Children and Families

**Guidelines for Expert Evaluations in
Child Abuse/Neglect Proceedings (Mental Health)**

November 8, 2012

Commissioner Allison Blake,

PhD LSW

**Child Abuse and Neglect Mental Health Evaluation and Treatment
Advisory Group Member List**

Chair: Debra Lancaster

Director
Office of Child and Family Health
NJ Department of Children and Families

Karen Baldoni

Director
Office of Contract Administration
NJ Department of Children and Families

Program Support Specialist
Office of Contract
Administration
NJ Department of Children and Families

Mary Beirne, MS, EdD, MD

Chief Child/Adolescent Psychiatrist
NJ Department of Children and Families

Lauren Carlton

Assistant Attorney
General NJ Division
of Law

Nancy Carre-Lee

Assistant Director
Division of Child Protection &
Permanency NJ Department of
Children and Families

Anthony V. D'Urso, Psy.D.

Section Chief and Supervising
Psychologist Audrey Hepburn
Children's House
Associate Professor - Graduate Clinical
Psychology Montclair State University

John Esmerado

Assistant Prosecutor
Union County Child Advocacy Center

Amanda Farley

Program Support Specialist
Office of Child and Family
Health
NJ Department of Children and Families

Jeffrey Guenzel

Deputy Commissioner
NJ Department of Children and Families

Kristin Huddy

Adrienne Jackson, MSW

Executive Coordinator
NJ Task Force on Child Abuse & Neglect

Walter Kalman, MSW, LSW

Executive Director
National Association of Social Workers, NJ
Chapter

Barry A. Katz, Ph.D.

NJ Psychological Association

Anthony P. Kearns, III

Prosecutor
Hunterdon County Prosecutor's Office

**Julie
Lippman,
Psy.D.**

Evaluation
Services
Director
CARES
Institute

Keri Logosso-Misurell, Esq.

Board Member
New Jersey Children's Alliance

Joseph Ribsam Jr., Esq.

Assistant Commissioner
NJ Department of Children and Families

Brian Ross, Esq.

Assistant Director
Office of Policy and Regulatory
Development NJ Department of
Children and Families

Aileen Torres, Ph.D.

Past-President
Latino Psychological Association of New
Jersey

Kara Wood

Director
Division of Child Protection
& Permanency NJ
Department of Children and
Families

I. Introduction

This is the Department's first comprehensive effort to address the use of expert evaluations in child welfare and child protective services proceedings. These guidelines lay out best practices for forensic evaluations and assessments that may be needed during child welfare investigations, to assist with permanency planning, or during litigation of guardianship complaints.

Child abuse and neglect cases are often complex. Expert consultants are frequently used to assist caseworkers, attorneys, law guardians, judges, and parents in making determinations, case planning, and decision making. The experts' services are often in the form of forensic evaluations of the mental health status and/or capabilities of the parents of dependent children. In addition, an evaluator may assess a child's behavioral functioning or developmental status as well. CP&P and the courts often rely on these evaluations and recommendations for effective case planning and to guide the court's decision making process.

In developing the guidelines that follow, the Department reviewed and analyzed professional guidelines and the work of other states, and convened an interdisciplinary group of experts to form DCF's Advisory Group on Child Abuse and Neglect Mental Health Evaluation and Treatment.

The role of the Advisory Group was to assist in formulating a framework that is flexible enough to accommodate differences in disciplines while providing clear practice guidelines that address the questions to be asked, the information required, the tools necessary to inform the evaluation, the credentials and qualifications of the evaluator, and the essential components of the evaluation itself.

The guidelines that follow are intended to improve the quality of expert forensic evaluations provided for CP&P and the courts, as well as the ability of stakeholders involved in child welfare proceedings and child protective service matters to make better use of them. It is clear that representatives of different disciplines with differing philosophical orientations will have varying approaches to the task of providing a forensic assessment. Each unique discipline will organize their work in a way that reflects their individual expertise. These guidelines are not meant to supplant the professional judgment of evaluators regarding their response to the unique features of each case.

The first sections of this document are general guidelines, followed by more specific recommended practices.

II. Definition/Application

For the purpose of these guidelines, a forensic evaluation in child welfare

proceedings and child protective service matters is an evaluation necessary to assist the court and/or CP&P in case planning, or to resolve a case. A forensic evaluation may be requested by CP&P, by another party to a proceeding, or the court. Any evaluation that may reasonably be expected to be submitted to the court is termed forensic. Although forensic evaluations may contain treatment recommendations, the primary function of the forensic evaluation is to inform the parties and to assist the court in rendering decisions in child welfare cases.

These guidelines do not cover evaluations or assessments obtained primarily for mental health treatment purposes, substance abuse, anger management, psycho-sexual evaluation, or domestic violence, although any or all of these issues may be addressed in a forensic evaluation.

These guidelines recognize that, in child welfare cases, the emphasis is on the safety, permanency, and well-being of the child.

III. General Principles and Guidelines¹

1. The Role and Function of Forensic Evaluations in Child Welfare Matters

The primary function of an evaluation is to provide a report that contains relevant, professionally sound observations, results and opinions in matters where a child's health and welfare may have been harmed or placed at risk of harm. To ensure the reliability of the evaluator's conclusions all opinions that are rendered must be given within a reasonable degree of medical/psychological/clinical certainty. The specific purposes of the evaluation generally will be determined by the referral questions and/or concerns provided to the evaluator by the referring party or parties. When the child already has been found by the court to be at risk of harm, the evaluation of the parent(s) generally identifies interventions intended to reduce future risk to the child, and often focuses on rehabilitation recommendations designed to protect the child and help the family. An additional purpose of such an evaluation may be to make recommendations for interventions that promote the psychological and physical well-being of the child, and, when appropriate, facilitate the safe reunification of the child with the parent.

Consistent with State law, evaluators appreciate the value of expediting family reunification, when possible and safe, while they also understand the value of other permanent plans when reunification is not possible.

The evaluation addresses the particular psychological, behavioral, and developmental needs of the child and/or parent(s). Relevant issues may include, but are not limited to, abuse or neglect of the child, safety, parental capabilities, or reunification or other permanency plans. In considering psychological factors affecting the health and welfare of the child, evaluators

may focus on caregiver capacities in the context of the psychological and developmental needs of the child. This may involve an assessment of:

- The adult's capacities for parenting, including those attributes, skills, strengths and abilities most relevant to abuse and/or neglect concerns;

¹ Washington State Supreme Court Commission on Children in Foster Care. "Guidelines for Expert Evaluations in Child Welfare Proceedings." Washington, 2007. Available online at, [http://www.courts.wa.gov/committee/pdf/Guidelines%20for%20Expert%20Evaluations%20in%20Child%20Welfare%20Proceedings%20\(2007\).pdf](http://www.courts.wa.gov/committee/pdf/Guidelines%20for%20Expert%20Evaluations%20in%20Child%20Welfare%20Proceedings%20(2007).pdf)

- The psychological functioning, behavioral, and developmental needs of the child, particularly with regard to vulnerabilities and special needs of the child, as well as the quality of the child's attachment to the parent(s) and the possible developmental and emotional effects of separation from the parent(s), siblings, extended family members, and other caregivers;
- The current and potential functional abilities of the parent(s) and, when necessary for resolution of the case, other relatives, to meet the needs of the child; and/or
- The need for and likelihood of success of clinical or other interventions for identified problems, which may include recommendations regarding treatment modalities and objectives, frequency of services, specialized interventions, parent education, and the child's placement.

2.

General Competencies of Expert Evaluators

Evaluators should gain and maintain specialized competence. Expert evaluators in child protection matters are aware that special competencies and knowledge are necessary for the undertaking of such evaluations. Competence in performing expert evaluations of children, adults and families is necessary but not sufficient. Education, training, experience and/or supervision in the areas of forensic practice, child and family development, child and adult psychopathology, the impact of separation on the child, the nature and consequences of different types of child abuse and neglect, and the significance of human differences may help to prepare evaluators to participate competently in expert evaluations in child protection matters.

Evaluators:

- Use current knowledge of scholarly and professional developments, consistent with generally accepted clinical and scientific practice, in selecting evaluation methods and procedures² and are aware of

evidence-based practices.

- Strive to become familiar with applicable legal and regulatory standards and procedures, including local State and Federal laws governing child protection issues. These may include laws and regulations addressing child abuse, neglect, and termination of parental rights.
- Describe the scientific basis for their judgments or recommendations, and state when their judgments or recommendations may expand on, or not be fully supported by, currently accepted clinical and scientific practice.
- Are aware of, and develop their knowledge and special competencies for, evaluation of specific populations including, but not limited to, issues related to literacy, the needs of persons who do not speak English, sensory impairment, psychological disorders, and developmental impairments.
- Should be fluent in the child's/parent's native language, when possible (have experience using a court appointed interpreter, if language presents a difficulty).

² Note: Examples of standard setting organizations include American Psychological Association, the National Association of Family and Conciliation Courts, The American Academy of Child and Adolescent Psychiatry and others.

- Have appropriate qualifications to conduct an evaluation and/or to testify at court, including language, cultural competency, and other qualifications specified in CP&P contracts.
- Should be competent in the cultural norms of the child/parent being evaluated.
- Utilize language and culturally correct testing.
- Have expertise in working with relevant clinical populations, including:
 - Children;
 - Sex offenders;
 - Domestic violence victims and batterers;
 - Persons with developmental disabilities; and,
 - Persons with psychiatric/neurological/neuropsychiatric diagnosis.
- Have expertise with the instruments employed, including psychological and intellectual tests that will need to be interpreted by a licensed psychologist, who is familiar with the norms and the uses of that test with the relevant population.
- Are experts in the use of appropriate interview techniques.
- Must not serve as an expert evaluator if they are the treating professional.

Evaluators must be aware of personal and societal biases and engage in nondiscriminatory practice. Evaluators engaging in expert evaluations in child protection matters consider how biases regarding age, gender, race, ethnicity, national origin, religion, sexual orientation, disability, language, culture, and socioeconomic status may interfere with an objective evaluation and

recommendations. Evaluators should be aware of the potential for defensiveness on the part of participants, given the circumstances, and must take this into account when conducting the evaluation and upon making recommendations. Evaluators recognize and strive to overcome any such biases. If unable to overcome his or her own biases, the evaluator will either withdraw from the evaluation or seek assistance in completing the evaluation. When interpreting evaluation results, evaluators must be aware that there are diverse cultural and community methods of child rearing, and consider these in the context of the existing local State and Federal laws. Also, evaluators should use, whenever available, tests validated with populations similar to those being evaluated.

Evaluators avoid multiple relationships to maintain objectivity. In conducting expert evaluations in child protective matters, evaluators avoid multiple role relationships. Evaluators generally do not conduct forensic evaluations in child protection matters in which they have provided clinical services for the child or the immediate family, or have had other involvement that may compromise their objectivity. Providing clinical services to the child or other participants following an expert evaluation is discouraged. A treating professional can be called to testify, but should NOT recommend a permanency plan.

3. Procedural Guidelines: Conducting an Evaluation

Evaluators and referring parties understand that forensic evaluations in child welfare and child protection matters may present a wide variety of legal and/or ethical considerations. Evaluators and all parties appreciate the need for timeliness in child protection matters (e.g., response to evaluation referral, scheduling appointments, completion of reports).

The purpose of the evaluation should be clear upon referral and should outline the specific questions to be addressed by the evaluation. In all cases, the referring party or parties should clearly state the purpose of the evaluation in writing and pose specific questions to be addressed in the evaluation.

Based on the nature of the referral issues and questions, the scope of the evaluation is determined in the referral or by the court, in consultation with the evaluator. Once the referral questions and scope of the evaluation have been accepted by the evaluator, the expert evaluator chooses appropriate methods with which to address the questions. Evaluators may also identify relevant issues not anticipated in the referral questions that could enlarge the scope of the evaluation; these should be conveyed to the parties as early as possible. For issues outside the scope of the evaluator's competency, the evaluator considers recommending additional services or evaluations.

Evaluators inform participants about the disclosure of information and the limits of confidentiality.

- When an evaluation is court ordered, it is not privileged and the evaluator informs the individuals of the nature of the evaluation and that the evaluation will be distributed to other parties as provided by court order. Evaluators conducting an evaluation in child protection matters ensure that the participants, including the child (to the extent feasible), are aware of the limits of confidentiality for the evaluation results. If the public agency or court is paying for the evaluation, the evaluator so informs the individual.
- When an evaluation is not court ordered, evaluators performing evaluations in child protection matters should obtain informed consent from all adult participants, and children and youth consistent with their developmental capacity to understand.
- When an evaluation is obtained by a party in an abuse/neglect or termination proceeding without the apparent knowledge or consent of the child welfare agency, guardian ad litem, and/or the court, the evaluator should advise the party being evaluated of the need to obtain and review appropriate and relevant information from the child welfare agency, guardian ad litem, and/or the court.

Evaluators use multiple methods of data gathering. Evaluators generally use multiple methods of data gathering, including, but not limited to, clinical interviews, observation, and/or psychological testing that are sufficient to provide appropriate substantiation for their findings. Evaluators should review relevant reports (e.g., information from child protection agencies, social service providers, law enforcement agencies, health care providers, child care providers, schools, and institutions). In evaluating parental capacity to care for a particular child or the quality of the parent-child interaction, evaluators should make reasonable efforts to perform formal observations of the child together with the parent, unless such observation is not necessary to respond to the questions posed in the evaluation or to support the recommendations and conclusions of the evaluator. Evaluators in some circumstances may rely on formal observations conducted by other neutral and competent professionals. It is recognized that in some circumstances, parent-child observations may not be necessary. Also, in some circumstances, it may not be advisable to require parent-child contact for purposes of the evaluation. For example, in cases where the safety or well-being of the child is clearly in jeopardy or parental contact with the child has been prohibited by the court. In such cases, the evaluator should note explicitly the reason(s) that a parent-child observation was not included. Evaluators may also interview extended family members and other individuals, when appropriate (e.g., caregivers, grandparents, and teachers). However, these should not be considered as substitutes for formal observation.

Evaluators are able to provide clarification and answer questions relating

to the evaluation(s) completed. Once an evaluation is completed, the evaluator must be available to speak with CP&P staff such as the assigned caseworker if there are any questions or concerns regarding the evaluation.

IV. Best Practices for Expert Forensic Evaluations

Forensic evaluations may be needed at any point in time during the lifespan of a child protective services case. The need for a forensic evaluation may emerge during the course of an investigation to assist with developing understanding or seeking clarity around the allegations of child abuse/neglect. More commonly, mental health evaluations may be required to contribute to the decisions by the court of the Division made about placement, reunification, permanency, and visitation. Finally, forensic evaluations are typically required for guardianship (termination of parental rights) litigation.

1. During an Investigation

The Role and Function of Forensic Evaluations during an Investigation

During an investigation, evaluations may be needed to assist CP&P and the Courts in assessing whether abuse and/or neglect occurred. These evaluations are meant to assist in clarifying or gathering additional information for investigative purposes with the lens of an expert. When sufficient evidence or clarity about the case has been achieved through the investigative work of the CP&P caseworker via interviews and collateral review, or teamed efforts with law enforcement or others involved in the investigative process, it is often not necessary to engage the services of an expert for an evaluation during an investigation

Forensic evaluations during the investigatory phase of the case may be warranted as part of the investigative efforts conducted by CP&P (and law enforcement). These situations most often include allegations of sexual abuse and emotional abuse/neglect. In addition, an evaluation during the initial involvement with a child may assist CP&P in determining the impact of an event on a child's psychological functioning.

Evaluations that may be required during the course of an investigation are almost always time sensitive matters. Thus, it is recommended that referrals be made as close to the point in time of the allegation or the occurrence of the alleged incident as possible:

- Evaluators should receive referrals within 10 working days of the report.
- An appointment by the evaluator should be granted within 10 working days of the referral.
- CP&P shall provide available background materials by the time of the evaluation.
- Evaluators should complete their reports and provide them to CP&P within 10 working days following completion of the evaluation.

These guidelines recommend that no more than 45 days pass between the initial referral to a provider for an assessment, to the date the written report, with recommendations, is provided to CP&P for review.

The Forensic Evaluation Process during an Investigation

In consultation with supervisory staff, and the DAG if litigation is contemplated or a complaint has been filed, CP&P caseworkers should select a provider who has the appropriate credentials to perform the evaluation. In many cases, child protection staff should access their Regional Diagnostic Treatment Center to conduct these evaluations. CP&P requires licensed individuals to conduct evaluations. In most cases, these will be licensed psychologists. When the impact of physiological factors, medical illness, medication, neurological, or psychiatric disorder is complex, an evaluation by a psychiatrist or physician may be necessary. In limited circumstances, an assessment by a LCSW may be appropriate.

The purpose of the evaluation during the investigatory phase of a case must be clear and should outline the specific questions to be addressed by the evaluator. Confirm with the evaluator the purpose of the evaluation. It is particularly important to limit the number of interviews or evaluations a child experiences for both validity reasons and to avoid re-traumatizing a child.

Investigation Evaluation Referral Questions:

- Is this child's presentation consistent with the allegation?
- To what degree has the child been harmed or traumatized by the event?
- Is this child able to participate in court proceedings?
- Other questions relevant to the specific case.

Evaluators should use multiple methods of data gathering.

The evaluator should be provided with certain background information, which includes:

- CP&P investigation report (or summary report) that is current/up to date;
- Existing prior psychological and psychiatric evaluations of the child and biological parent(s);

- Available law enforcement records including police reports; criminal charges and convictions; taped interviews, if available; and Promis/Gavel history of offenses;
- Prior CP&P history, including all prior referrals, with a finding for each allegation/investigation; investigative summaries;
- Complaint filed in court; and,
- Known mental health, substance abuse, or domestic violence history.

If a child is to be evaluated, the CP&P caseworker assigned to the case should accompany the child to the evaluation to support the child, to be available to provide any additional information and to hear directly from the evaluator any initial findings or recommendations. Whenever possible so as to best inform the evaluation, the investigative worker should accompany the child. Whenever possible a trusted adult should also accompany the child.

During the clinical interview, an evaluator:

- Establishes “Ground Rules” between the evaluator and the child.
- Explains to the child, in age appropriate and developmentally appropriate terms, the nature and the scope of the evaluation.
- Establishes the child’s developmental and cognitive ability to participate in the evaluation.
- Establishes the child’s competency. Does the child know the difference between the truth and a lie, real or pretend?
- Obtains the child’s version of the incident.
- Notes the child’s affect upon describing the incident.
- Asks questions to gather past history.
- Determines family relationships.
- Determines peer relationships.

Once the evaluation has been completed, the summary and report should include:

- Reason for the report – summary background;
- Nature of the allegation;
- Prior history;
- Documentation including a summary of the interview and direct quotes by the person being interviewed;
- Clinical finding and explanation;
- Any formal diagnosis;
- Clinical determination – indicate whether supported/not supported; and,
- Recommendations.

2. During Permanency Planning/Hearings

The Role and Function of Forensic Evaluations during Permanency Planning/Hearings

At the time of referral and over the course of a child or adolescent's time under CP&P custody, mental health evaluations may be required to contribute to the decisions made about placement, permanency, and parental rights. During permanency planning and hearings, evaluations are often used for:

- Interim Evaluation: The interim evaluation is meant to guide reasonable efforts for reunification.
- Ten Month Conference: The ten month conference is used to prepare a permanency plan for the child or youth in out-of-home placement. Before moving forward, any previous reports should be reviewed. It would be useful if the evaluator from the interim evaluation was also utilized at this point.
- Periodic Evaluation – Evaluation of Imminent Concerns Arising during Placement: An evaluation of imminent concerns is used to assess any risks or challenges that the child may incur during the course of the protective services or guardianship litigation. Examples include:
 - Disruption of the current placement;
 - Acute crisis (e.g., psychiatric hospitalization, severe medical illness, runaway, arrest, school disruption); and
 - Significant change in response to visitation.

Forensic Evaluation Process during Permanency Planning/Hearings

The purpose of the evaluation should be clear and should outline the specific questions to be addressed by the evaluation. The following referral questions should help to guide forensic evaluations at each of the stages identified for permanency planning/hearings:

Interim Evaluations Referral Questions:

- What services are needed for reunification?
- What impact has the abuse/neglect history had on the child?
- What are the risks that need to be addressed?
- Is the parent fit and able to parent the child?
- What actions are recommended to address the risks?
- What are the strengths that can be built upon?
- What visitation can be safely afforded between parents and their child(ren)?

Ten Month Conference Referral Questions:

- What progress has been made towards eliminating the harm?
- What still needs to be done?
- Are there any new areas of need?
- If a home other than the child(ren)'s current placement is being considered, is it in the best interest of the child(ren) to move to another placement if proposed by the parents, or to stay permanently where he or she is residing?
- Can this child transition back to the biological parents, without experiencing more harm than good?
 - If bonding and attachment are issues, an evaluation by a psychologist is necessary. A psychiatrist may contribute information within his or her area of expertise.
- Have the correct services been provided so far, and is there a need for a reduction, modification, or expansion of services?

It may be necessary to reevaluate the permanency plan. All of the questions above would apply to any such reevaluations.

Evaluators should use multiple methods of data gathering.

For these evaluations, the evaluator should be provided with certain background information, which includes:

- Existing prior psychological and psychiatric evaluations of the child and biological parent(s);
- Existing treatment reports for biological parents and child;
- Known mental health, substance abuse, or domestic violence history;
- Visitation reports;
- Complaint for guardianship, if filed;
- CP&P investigation report (or summary report) that is current/up to date;
- Prior CP&P history, including all prior referrals, with a finding for each allegation/investigation; investigative summaries;
- Most recent CP&P court report;
- Important selected contact sheets from the CP&P case record;
- Available law enforcement records including police reports; criminal charges and convictions; taped interviews, if available; and Promis/Gavel history of offenses;
- Additional information the parent wants to share with the evaluator; and,
- Any further available information requested by the evaluator.

All evaluations should include a review of comprehensive, accurate background information; a clinical interview; and the use of an appropriate assessment tool.

The evaluator should have access to all information he or she deems necessary in order to respond to the questions posed.

Periodic Evaluation – Evaluation of Imminent Concerns Arising during Placement Referral Questions:

- Identify impact of presenting problem.
- What are the recommended services or actions to address the problem?
- Should the permanency plan change?

For Periodic Evaluations of Imminent Concerns Arising during Placement, documented relevant information is needed as well as all available relevant reports, such as:

- Medical reports;
- Police reports;
- School reports;
- Psychiatric reports; and
- Relevant contact sheets.

3. During Litigation for Guardianship Complaints

The Role and Function of Forensic Evaluations during Litigation for Guardianship Complaints

Guardianship evaluations consist of fitness and bonding assessments during trial preparation after a guardianship complaint has been filed. Ideally, both the fitness and bonding assessments are completed by the same psychologist.

The presumption is that fitness and bonding assessments are required for guardianship litigation. It is recognized that in some circumstances, parent-child observations may not be necessary or advisable for purposes of the evaluation. For example, in cases where the safety or well-being of the child is clearly in jeopardy or parental contact with the child has been prohibited by a prior fitness and bonding assessment, parent-child observations may be bypassed. In such cases, the evaluator should note explicitly the reason(s) that a parent-child observation was not included.

A bonding evaluation assesses the relationship between the child(ren) and the proposed caregivers and other household members as appropriate.

Forensic Evaluation Process during Litigation of Guardianship Complaints

Guardianship Evaluation Referral Questions:

- What progress has been made towards eliminating the harm?
- What still needs to be done?
- Are there any new areas of need?
- If a home other than the child(ren)'s current placement is being considered, is it in the best interest of the child(ren) to move to another placement if proposed by the parents, or to stay permanently where he or she is residing?
- Can this child transition back to the biological parents, without experiencing more harm than good?
 - If bonding and attachment are issues, a psychological evaluation is necessary. A psychiatrist may contribute information within their area of expertise.
- Assess the child's bond and attachment to the biological parent(s).
- What harm, if any, will result if parental rights are terminated?
 - Can the resource family parents mitigate the harm?
- Assess the child's bond and attachment to any proposed adoptive resource parent(s).
- Would severe and enduring harm occur if the child is removed from the proposed adoptive resource parents?
 - Can the biological parents mitigate the harm?

Guardianship evaluations call for specific competencies³ that are referred to in this section.

The evaluator at this stage in most circumstances will be a *licensed psychologist* or a psychiatrist. The licensed professional must be qualified to perform custody/parenting time evaluations and/or termination of parental rights evaluations through education, training, and/or supervision in all of the following categories:

1. Child growth and development;
2. Psychological testing;
3. Parent-child bonding;
4. Parenting skills;
5. Adult development and psychopathology;
6. Family functioning;
7. Child and family development;
8. Child and family psychopathology;
9. The impact of divorce or family dissolution on children; and,
10. The impact of age, gender, race, ethnicity, national origin, language, culture, religion, sexual orientation/identity, disability, and socioeconomic status on custody/parenting time evaluations.

When the following topics are involved, the licensed psychologist or psychiatrist shall have specialized education, training, and/or supervision in the specific topic, or the licensee shall refer to a licensed mental health care provider who has that

education, experience, training, and/or supervision. The topic areas include:

1. Physical, sexual, or psychological abuse of spouse or children;
2. Physical and emotional neglect of children;
3. Alcohol or substance abuse that impairs the ability to parent;
4. Medical/physical/neurological impairment that affects the ability to parent; or
5. Other areas beyond the licensee's expertise that are relevant to the custody/parenting time evaluation.

Evaluators may identify relevant issues not anticipated in the referral questions that could enlarge the scope of the evaluation. At this stage, it is important to consider some relevant factors or issues in responding to the bonding and attachment referral questions.

These factors include:

1. Age of the child;
2. The developmental stage of the child;
3. Child's history of abuse and/or neglect;
4. Child's resiliency;
5. Any special needs - medical or emotional - of the child or biological parents;
6. Parenting skills of both sets of parents;
7. Length of time in biological parents' care;
8. Number of placements;
9. Length of time in each placement;
10. Previous failed reunification attempts;
11. Child's wishes, weighted in accordance with developmental functioning;
12. Demonstrated willingness and ability of both biological parents and proposed adoptive resource parents to comply with services;
13. Demonstrated willingness and ability of both biological parents and proposed adoptive resource parents to recognize and meet the child's needs, including issues relating to reunification or adoption;
14. History of child's interaction with both biological parents and proposed adoptive resource parents;
15. Issues that may affect child's behavior during a bonding evaluation; and,
16. Sibling bonds/other attachments.

Evaluators should use multiple methods of data gathering.

Evaluators should be provided with the same background information listed under *Section 2: During Permanency Planning/Hearings*.

All evaluations should include a review of comprehensive, accurate background information; a clinical interview; and the use of an appropriate assessment tool. The evaluator should have access to all information he or she deems Necessary in order to respond to the questions posed.