All Unhappy Families: Standardization and Child Welfare Decision-Making

by

Emily Adlin Bosk

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Social Work and Sociology)
in the University of Michigan
2015

Doctoral Committee:

Professor Karin A. Martin, Co-Chair
Professor Mary C. Ruffolo, Co-Chair
Associate Professor Renee Anspach
Associate Professor Beth Glover Reed

This dissertation is dedicated to the memory of
Lesley Siegel
who worked ceaselessly to improve
the mental health of our most vulnerable children.

Acknowledgements

When I was stuck, I would imagine writing these acknowledgements. Calling up the wisdom, generosity, and support that so many people have furnished, in so many different ways, at so many different times, always made it possible to find a way through. I have many, many, many people to whom I am indebted:

Karin Martin offered unparalleled guidance and insight throughout every stage of the dissertation process. One conversation we had best illustrates her seemingly effortless ability to improve, clarify, and sharpen. When I called her from the field with heavy traffic making it difficult to hear, Karin took some jumbled thoughts and transformed them into a clearly articulated question that has guided this research ever since. Throughout graduate school, Karin has routinely sparked my sociological imagination with her dedication to incorporating children's voices and experiences into spaces from which they have been absent. Her support has also been personal. Karin has made room for chronic illness and motherhood, insisting that I can be part of the academic community and a person at the same time.

Mary Ruffolo is an academic Jedi. I have never seen anyone quite so good at navigating and shaping institutions. As I progressed through the program she gave me her unwavering support and enthusiasm, pushing me to integrate the theoretical with the applied. Mary strategized, problem-solved, and thought deeply about the whole scope of this project more times than I can count. Ahead of the curve, Mary already promotes reflective systems. Like Karin, her support has been personal as she too, cleared a space for me to balance life within the academy

Not many people have mentors who brought them dinner and visited after the births of their children but Mary and Karin did. They are my models for what it

means to be a professor. I admire them both immensely.

Renee Anspach was the very first person I spoke with at the University of Michigan and her introduction made it obvious that this was the place to learn and grown. Renee's brilliance was on display again and again through draft after draft. She caught weaknesses in arguments no one else saw and patiently offered her time when it was in short supply. By helping me to connect this work to organizational theory and medical sociology; Renee expanded its scope. The care and precision of Renee's own scholarship will always offer a template for my own.

Beth Glover-Reed has championed my interests from the beginning of my time in the joint doctoral program. Her commitment to intersectional thinking, courage in engaging complexity in research, and contention that research should mean something, push me to do the same. Beth saw early on that this study should be framed around the Structured Decision-Making tools and helped me to tease out their role as boundary objects. Beth's voice will always be in the back of my head as I work to connect my research to issues related to social justice.

Sandy Danzinger, Julie Ribaudo, and Karen Staller opened their doors for long conversations that often shifted or crystalized my thinking and encouraged my efforts when they had no formal obligation to do so. In their own ways, each of these women contributed to helping me locate myself in the field. I am grateful for their always sage advice and friendship throughout the joint doctoral program. As graduate directors, Berit Ingersoll-Dayton and Sarah Burgard supported me and my work especially through the interruptions brought by motherhood.

Lesley Siegel, to whom this dissertation is dedicated, was my policy mentor through the Doris Duke Fellowship. I met Lesley when she taught a seminar at the Yale Child Study Center and was quickly impressed, like so many who knew her, by the ways she interwove clinical and systems level thinking. Lesley always questioned whether she had the experience to provide mentorship throughout this research. Yet, she implicitly understood the points I was trying to make before I was

able to articulate them myself. Lesley made a comparative project possible, using her connections to open doors to one state's child welfare system. Her insight and encouragement were foundational to this research and her commitment to addressing issues of mental health and trauma within the child welfare system will continue to guide me throughout my career. I am grateful for the conversations we had and miss the ones we will never have.

There are people who I cannot thank publicly without jeopardizing my commitment to participant confidentiality. I am deeply grateful to the four agency directors in both states who gave their blessings to this work and access to their staffs. They do the hard, thankless tasks, of serving families and working to improve the child welfare system every day. Thank you, thank you, thank you.

At the Child Advocacy Law Clinic at the University of Michigan, Vivek Sankaran, Josh Kay, and Kate Rosenblum patiently and sensitively untangle the complexities of the child welfare system semester after semester. They have shown me how its done. Vivek provided not only mentorship but also key introductions, without which I would not have been able to conduct this research.

The 66 people who I interviewed for this study shared their experiences and time. There is nothing easy about being a front line child welfare worker. The cases are hard and sometimes heartbreaking, the resources are limited, the stakes are high, and the countervailing pressures are many. Child welfare workers are often blamed when things go wrong with little attention paid to the structural difficulties in which they operate. I appreciate the trust each participant instilled in me when they sat down to tell me about what they do and how they do it. Although their stories are represented only in snapshots throughout this dissertation, I hope that I have done justice to the multi-layered nature of their work and the many complexities that surround it.

Judy and Stocky Clark opened their home up to me for a second time in as many years while I collected data. Judy cooked me dinner after long days in the field,

refused to confuse jargon with good thinking, and didn't seem all that fazed when I couldn't tell her when I would be leaving…

This work was supported by grants and fellowships that provided critical time, space, and money to see it to its completion. The National Science Foundation Law and Social Science Dissertation Research Improvement Grant, the Fahs-Beck Fund for Research and Experimentation, the Rackham Centennial Award and the Vassar Graduate Fellowship Awards were essential to undertaking a project that often took me far from home. The Doris Duke Fellowship in the Prevention of Child Maltreatment and the Promotion of Child Wellbeing not only offered two years of support but also created an interdisciplinary network of early career scholars that I cherish being a part of. I am grateful for the many opportunities that accompany being part of the program. Internal grants and awards from the University of Michigan, the Rackham Graduate School and the Joint Doctoral Program: were instrumental to my growth and development as a scholar and gave me confidence that I had something to say.

Pat Preston offered unparalleled expertise in navigating the ins and outs of applying for and administrating external funding. While Morgan Purrier and Christopher Luoma provided enthusiastic and excellent research assistance.

The Soc of Awesome Group unironically lived up to its tongue in cheek moniker. Dan Hirschman, Elizabeth Young, Jonah Siegel, Kathy, Lin, Michelle Phelps and Rachel Pierotti were partners in commiseration, anxiety management, accountability, writing and revision. They offered constructive criticism on the prospectus, talked through all the ins and outs of crafting a dissertation, and thought through my analytic strategy and emerging analyses. Dan Hirschman, who shares my enthusiasm for deconstructing processes of commensuration, stood in a hallway on more than one occasion to puzzle through something, even when it was clearly inconvenient to do so. Elizabeth Young made coming into the office something to look forward to. I thank each of them for their incredible intellectual generosity and friendship.

Before I came to Michigan, I was told that graduate school would be a lonely and isolating experience. I gulped but forged ahead anyway. Thank god I did. If I hadn't, I would not have met some of the best people that I have ever known. Francesca Forzani and Jeremy Sussman, Jessica Wyse and Stephan Linder, Beth Schwanke, Adrienne Fowler, Isabel Daniels, Elizabeth Armstrong, and Maria Wathen lived life with me through all the changes that happen over almost a decade of training and offered the profound gift of constant friendship as graduate school and dissertating unfolded.

It occurs to me that after many years, I know very little about some of my friends' work. Instead, I have the privilege and the pleasure of knowing that this gang will freely offer to move 4 couches in as many weeks, make you laugh until it hurts to breathe, and that many, many people can be fit into 1200 Iroquois for football watching, friends-giving celebrating, and general chicanery. Alex Jakle, Claire Whitlinger, Alex Von Hagen-Jamar, Kerri Leyda Nicoll and Chris Leyda, Betsy and John Bringewatte, Lisa and Matt Watkins, Molly Reynolds, Joel Ruhter, Tim Ryan, Allison Dale-Riddle and Matt Riddle, Ashley Jardina, Dave Cottrell and Alton Worthington, *Candidates* will always be my favorite show.

When I first met Tova Walsh, I knew she was a kindred spirit. Wise, generous, fun, and funny, I owe so much of everything about the process and products of the joint program to my dear friend and colleague. Everyone should be so lucky to have a Tova in their lives. Believe me, I know I am.

My father in-law, Morton Schoolman, is the kind of academic who reads Proust on the beach. His passion for his vocation and dedication to his students and colleagues are a model. My mother in-law, Maureen Schoolman is a magnanimous and talented copy editor. Over the years she has seen this project in all its guises and saved me from more than one embarrassing grammatical error and poorly expressed thought. I didn't have the heart to make her read the complete version, so any misplaced comma is my fault alone.

Abbie Bosk kept me going with perfectly timed texts, cheers from afar, and sole vacations spent with her nephews helping out. A pediatrician devoted to improving care for inner city kids, it's still not clear who wins our ongoing contest to quiet dinner parties just by talking about our work. What is clear is that my sister is the best ever and my BFF.

Many people say their parents are their inspirations, but Marjorie and Charles Bosk really, really, really are. A brilliant psychologist and sociologist, they taught me to see the world in all its social and individual complexities. My father never talked to me like I was child; instead he gave the singular gift (among many) of sharing his grand ideas and encouraging me to come up with my own. My mother taught me that all of life is about making meaning and that this is best done in the company of others. As first generation everythings, my parents expected my sister and I to pay attention to all the ways in which the world is unfair and unjust and then they showed us how we could go about trying to make it better. They sacrificed, saved, and sweated to give us the privileges that they did not have. These privileges are incommensurable and innumerable. But I can and will quantify this: the greatest one of all is being their daughter.

My babies, Milo Benjamin and Finn Alexander were born in the middle of this dissertation. They remind me daily of my good fortune and their gentle, yet feisty, spirits make both the best and worst days of writing seem insignificant.

If you're going to be in a foxhole, you want to be in there with Ethan David Schoolman. We have been in one too many together in the time that this dissertation was written and this work just simply couldn't and wouldn't have been without him. Ethan's humor, support, grace under fire, and ability to work through anything are some of the many qualities that I admire most about him. His commitment to an equal partnership and care for our children are a clear example that the world can and does change. Ethan gives me hope about everything. He has sustained me, supported me, challenged me, encouraged me and loved me when I've

been absolutely impossible.

I'm grateful for it all, but most especially for him.

Table of Contents

Chapter 1

Introduction

Child Welfare Decision-Making in Context

Child welfare decision-making is a high-stakes task. When decisions about whether child maltreatment has taken place are incorrect; children left in homes who should have been removed can die or experience further abuse. Separations between children and their primary caregivers can cause lasting disruptions in their social, emotional, and cognitive development (Bowlby, 1969/1982, 1979, 1980; Wilkes, 1992) as well as be deeply traumatic to caregivers, especially when they are unwarranted. Once children have been removed from their families, the simple fact of the separation itself can make reunification difficult. Figuring out how to do decision-making well is just as highly fraught as the actual decision-making itself. Silver bullets and new policies wrapped in powerful epistemological frameworks come in and out of fashion, capturing and then losing, the attention of researchers, advocates, and policymakers when they fail to be the corrective originally claimed. At the center, are children and families whose lives often hinge on how decision-making about them is collectively understood and organized.

As a home-visiting social worker working with families with children at risk for removal and as a psychological consultant at the Child Advocacy Law Clinic at

the University of Michigan, I experienced first hand how difficult it is to make sound decisions about whether maltreatment is taking place, how severe the maltreatment is, if it is indeed occurring, and what to do about it. What is obvious on paper and in theory, is perpetually messy, inchoate, and ambiguous in practice. Navigating the complex realities of intersecting social issues: poverty, a failing mental health system, a troubled economy with few employment opportunities for those on the lowest rungs of the socio-economic ladder, and limited long-term treatment options for families makes decision-making about already complicated cases even harder. With so few modes of intervention that seem capable of addressing the systemic issues that have brought a family to the attention of Child Protective Services (CPS), what to do and how to do it is almost never clear.

In my work, I have witnessed what happens when mistakes are made and have made some of those mistakes myself. I have also worked with many dedicated people across professions who are committed to making the best decisions that they can for children and families. Yet, I am often left with the sense that we are collectively failing to do our job as well as we can. Qualitative research is guided by sensitizing concepts that are deployed as orienting frameworks (Blumer, 1979) 1979). Such concepts are starting points for a grounded theory (Corbin & Strauss, 2007) The sensitizing concepts motivating my dissertation research are drawn from my experience in the field and shape the outlines of this work.

In the past 15 years, the field of child welfare has seen a revolution in the technologies and practices related to how decisions are made. Actuarial-based assessments, now in use in the majority of states, are designed to make child welfare decision-making more scientific, fair, effective, and cost-efficient (CRC, 2008). These

are important goals when the weaknesses of clinical decision-making are well known, when decades of racial, ethnic, and socio-economic disproportionality challenge all claims to a system that is equitable and when research has consistently highlighted the inconsistency of decisions themselves (Dettlaff et al., 2011; Dettlaff & Rycraft, 2010; Fluke, Yuan, Hedderson, & Curtis, 2003; Gambrill & Shlonsky, 2000; Lindsey, 1992; Lu et al., 2004; Rossi, Schuerman, & Budde, 1996; Schuerman, Rossi, & Budde, 1999; Wulczyn & Lery, 2007).

Despite the popularity of actuarial-based risk assessments, we know relatively little about how these new technologies are playing out on the ground. Research about actuarial-based risk assessments has predominantly been through validation studies, which have lauded the ability of actuarial-based assessments to improve decision-making. Over 100 studies have shown that empirical decision-making models are superior to clinical judgment (Baird & Wagner, 2000; Baird, Wagner, Healy, & Johnson, 1999; Grove & Meehl, 1996). Studies specifically about actuarial-based risk assessments in child welfare are optimistic, overall, about their capacity to facilitate better decisions (Baird & Wagner, 2000; Baird, et al., 1999). What concerns there are about actuarial-based risk assessments in the child welfare community are often not about the concept itself but either methodological (e.g. appropriateness of the measures risk assessments employ) or practical (e.g. overconfidence in their ability to make perfect decisions by those who use them) in nature (Camasso & Jagannathan, 2000; Gambrill & Shlonsky, 2000, 2001; Munro, 2004; Wald & Woolverton, 1990).

While structured decision-making models have not explicitly been the targets of critiques themselves, larger challenges to the Evidence-Based Practice

Movement of which actuarial-based risk assessments are a part, apply. This set of scholarship raises questions about the consequences of removing contextual understanding and practice-wisdom from social work interventions and challenges the prudence of translating a model that began in medicine to the very different kinds of problems that social work addresses (Staller, 2006; S. Witkin & Harrison, 2001; S. L. Witkin, 1996).

Beyond theoretical debates about the strengths and weaknesses of Evidence-Based Practice in social work, research in implementation science has demonstrated that any intervention is only as valid and reliable as its deployment on the ground (Durlak & DuPre, 2008; NIH, 2013; Powell et al., 2015). Yet we know almost nothing about how actuarial-based decision-making in child welfare has been implemented. To date, there are only three qualitative investigations of actuarial-based risk assessments as they are used in practice and they raise serious questions about whether workers use these tools as intended. Lyle and Graham (2000) found that workers in Illinois inflated risk scores to obtain services for their clients, transforming the actuarial-based risk assessment from a neutral process for decision-making into a tool to advocate for clients (Lyle & Graham, 2000). The second found that hospital risk assessments were ignored when psychiatrists made decisions about which psychiatric patients to transfer and discharge from a maximum security hospital in favor of the information gathered during clinical interviews (Hilton & Simmons, 2001). A third study of court counselors revealed that they did not utilize their risk assessment when making probation decisions (Krysik & LeCroy, 2002). As a whole, these studies are not encouraging about the ability of actuarial-risk assessments to address the very issues that they are designed

to conquer and highlight the need for more sustained work on the role that standardized decision-making models have on decision-making in practice.

The very limited qualitative work on actuarial-based risk assessments raise more questions then they answer: (1) Why do workers ignore risk assessments? (2) Under what conditions to they ignore them? (3) When do workers use them as intended? (4) How does office policy intersect with workers actions and inactions regarding actuarial-based risk assessments? And (5) what is on the actuarial-based risk assessment itself that might influence a workers response to it? While the research cited above suggests workers to be homogenous in their responses to actuarial-based models, sociological theory would expect us to see heterogeneity in both worker and organizational attitudes about, and approaches to, actuarial-based risk assessments. Leading me to ask: (6) Are there variations among worker and organizational responses to risk assessments? (7) If such variation exists, how do we understand it? It is in this context that I ask the central questions that inform this dissertation: How do structured decision-making processes play out on the ground? What are the intended and unintended consequences of using them to frame decision-making?

My Approach

This dissertation applies a sociological lens to understanding how child-welfare decision-making is practiced in the context of standardized tools. In the three articles that comprise this work, I extend existing scholarship in social work on actuarial-based risk assessments to foreground how child welfare workers and the organizations in which they conduct their work, make-meaning of decision-making in a structured environment. Understanding the meaning-making processes

provides insight into what becomes salient as workers conduct their investigations and how larger organizational processes impact these considerations. Looking at the implementation of actuarial-based assessments at multiple levels also allows for a complex reading of the different variables that shape the adaptation of an intervention on the ground.

"According to Mantzoukas (2004), it is essential that researchers conducting qualitative research begin their work by presenting their epistemological stance, demonstrating "the rules by which they have agreed to play" (Bringewatt, 2013; Mantzoukas, 2004). As a sociologist, I do not assume that an actuarial-based approach or a clinical one is better. Rather, I take a constructionist and constructivist stance, focused on understanding how both individuals and organizations make meaning of the actuarial-risk assessment through every day work routines and through the creation of organizational practices and policies (Crotty, 1998). I view individual and collective responses to the actuarial-based risk assessment to be dialectical in nature, shaping and changing each other. Individuals must negotiate and square their views of the world with the organizational rules they are asked to follow and organizations must manage the individuals that comprise them. I understand the actuarial-based risk assessment itself to be both a scientific tool and a socially constructed practice whose dual technical and ontological nature must be interrogated. In this way, this dissertation follows in a long tradition of ethno-methodological sociology, which seeks "to discover the things that persons in particular situations do, the methods they use, to create the patterned orderliness of social life"(Garfinkel, 1984).

Debates about the 'correct' way to practice and the utility of actuarial vs. clinical approaches have been at the center of social work for the last 20 years. At their core, they are about nothing less than what it means to be a social worker, the values the profession holds, and the legitimacy of the field. These debates have played out with the passion and vigor one would expect of discussions that touch on such foundational concerns. By using sociological theories and methods to understand how individuals and organizations use actuarial-based risk assessments, my goal is, in part, to understand how these distinct and powerful frames are negotiated on the ground.

Given the primacy of both epistemologies and the intensity that discussions about each one engenders, one would expect a certain amount of confusion and noise as these opposing ideas are operationalized by social workers in the field. Throughout the dissertation, I have attempted to reframe actuarial and clinical epistemologies from mutually exclusive to co-existing and co-occurring with the goal of demonstrating how these views of social work practice are opposed, reconciled, or ignored (sometimes all at the same time) in the every day performance of social work by child welfare workers.

The Dissertation

The first paper from my dissertation investigates how and to what extent CPS workers have incorporated standardized methods like the Structured Decision Making Model (SDM) into their everyday work practices and then analyzes whether standardized assessments control for bias and subjectivity or encode bias and subjectivity in new ways. My research demonstrates that workers' ways of utilizing the SDM are associated with their socio-demographic characteristics, including

those related to gender, race, and years of experience. White male workers and workers with over 8 years of experience reported that they adjusted the SDM to ensure that its outcome reflected a decision they had already made. In contrast, African-American, female, and relatively inexperienced workers reported that they relied on the SDM to tell them what to do in cases and "did not make decisions." Contributing to a larger debate about the utility or limitations of actuarial risk models for the reduction of bias in decision-making, these findings contradict claims that SDM tools are neutral objects that facilitate objective conclusions about which families should maintain custody of their children. This work challenges the idea that standardized tools are able to eliminate subjective processes in decision-making by demonstrating that standardization alone cannot correct for differences in social power and standing that are embedded in any social system.

In response to findings like the ones describe above, disparities in child welfare outcomes are often investigated and then addressed at the level of the individual worker. Implementation science and Quality Improvement (QI) initiatives recommend better training to improve workers' 'fidelity to the model,' on the theory that the discrepancies between the desired and actual outcomes of an intervention are the result of workers' poor implementation. The second paper of my dissertation demonstrates that differences in decision-making are not confined to individuals but also occur at the organizational level. Comparing work practices in two states, I find that there are significant policy differences related to the meaning, implementation, and use of the SDM. These differences suggest that whether a standardized tool achieves a standardizing function is not only a consequence of the validity or reliability of the tool itself but also of the social and structural context in which a

tool is introduced and utilized. By re-framing standardization as a social rather than technical process, this paper challenges the idea that well-implemented, evidenced-based interventions with close fidelity to the model will necessarily be able to eliminate the messiness inherent in the complex work of child welfare decision-making. In doing so, I show that both tools and evidence-based practices are not self-actuating. My findings here have implications not only for Quality Improvement initiatives in social work practice but also for standardized tool-based approaches in other domains, such as education, medicine, and law.

In the third paper, I examine the construction of one state's actuarial-based risk assessment. In attempting to avoid errors where children are left in the homes when they should have been removed, I argue that both the risk assessment and the policies surrounding its use may set the stage for more unwarranted separations and punitive consequences for families. Using sociological theories of commensuration to examine how a 'trust in numbers' have come to organize how cases are evaluated and sorted, this paper closely examines what has been selected to be a part of the risk assessment and what gets left out (Porter, 1995). My research reveals that an explicit focus on demographic characteristics such as the number of children in the home, the age of the children, the number of past investigations and whether the parents were abused as children weight the risk assessment heavily in the direction of classifying a case as high-risk before the dynamics between the child and parent or the present circumstances have been considered. I argue that while the risk factors cited above have been well validated in the literature, in practice, their presentation on the risk assessment removes a contextual and interactional form of assessment from the evaluation process. State policies that prevent workers from

overriding risk assessment scores to de-escalate cases, make clinical judgment suspect in cases where risk may have been overestimated. Conversely, state policies that encourage, and at times, require workers to override risk assessment scores emphasize the importance of the ecological context for cases where risk may have been underestimated.

Throughout each paper, I take up the broader policy and practice implications of this research. Breaking away from rigid dichotomies that pit clinical and actuarial decision-making as mutually exclusive methods, I propose an integrative model that capitalizes on the strengths of both clinical and standardized processes. I argue that it is critical to understand and recognize how reflective practices and subjective assessments are incorporated into the processes for using standardized decision-making tools. Acknowledging these practices in actuarial processes whether in the form of narrative recommendations, reflective supervision about use of the tools in each case, or clinical case reviews, creates a space for unconscious, subjective, or biased assessments to play out. I find that when workers' subjective assessments are intentionally excluded from case decision-making processes, they find alternative ways to influence case outcomes regardless of the formal procedural mechanisms in place. The solution is not to find new ways to exclude caseworkers' judgment but to offer a place for it alongside other approaches to improving case decision-making. In the dissertation, I propose revising questions on the risk assessment to highlight relational and dynamic processes and suggest implementing structural opportunities for clinical review.

Works Cited

Baird, C., & Wagner, D. (2000). The relative validity of actuarial and consensus based risk assessment systems. *Children and Youth Services Review, 22*, 839-871.

Baird, C., Wagner, D., Healy, T., & Johnson, K. (1999). Risk assessment in child protective services: consensus and actuarial model reliability *Child Welfare, 78*(6), 723-748.

Blumer, M. (1979). Concepts in the Analysis of Qualitative Data. *The Sociological Review, 27*. doi: 10.1111/j.1467-954X.1979.tb00354.x

Bowlby, J. (1969/1982). *Attachment and Loss: Attachment* (Vol. 1). New York: Basic Books.

Bowlby, J. (1979). *The making and breaking of affectional bonds*. New York: Basic Books.

Bowlby, J. (1980). *Attachment and Loss: Loss Sadness and Depression* (Vol. 3). New York: Basic Books.

Bringewatt, E. (2013). *Making Sense of the Label: What Emerging Adults Can Tell Us About Growing Up with a Mental Health Diagnosis*. Ph.D., University of Michigan, Ann Arbor.

Camasso, M. J., & Jagannathan, R. (2000). Modeling the reliability and predictive validity of risk assessment in child protective services. *Children and Youth Services Review, 22*(11-12), 873-896.

Corbin, J., & Strauss, A. (2007). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Los Angeles: Sage.

CRC. (2008). The Structured Decision-Making Model: An Evidenced-Based Approach to Human Services: Children's Research Center-A division of the National Council on Crime and Delinquency.

Crotty, M. (1998). *The Foundations of Social Research: Meaning and Perspective in the Research Process*. London, England: Sage Publications.

Dettlaff, A. J., Rivaux, S. L., Baumann, D. J., Fluke, J. D., Rycraft, J. R., & James, J. (2011). Disentangling substantiation: The influence of race, income, and risk on the substantiation decision in child welfare. *Children and Youth Services Review, 33*(9), 1630-1637. doi: 10.1016/j.childyouth.2011.04.005

Dettlaff, A. J., & Rycraft, J. R. (2010). Factors Contributing to Disproportionality in the Child Welfare System: Views from the Legal Community. *Social work, 55*(3), 213-224.

Durlak, J., & DuPre, E. (2008). Implemenation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350.

Fluke, J. D., Yuan, Y. Y. T., Hedderson, J., & Curtis, P. A. (2003). Disproportionate representation of race and ethnicity in child maltreatment: Investigation and victimization. *Children and Youth Services Review, 25*(5-6), 359-373. doi: 10.1016/s0190-7409(03)00026-4

Gambrill, E., & Shlonsky, A. (2000). risk assessment in context. *Children and Youth Services Review, 22*(11-12), 813-837.

Gambrill, E., & Shlonsky, A. (2001). The need for comprehensive risk management systems in child welfare. *Children and Youth Services Review, 23*(1), 79-107.
Garfinkel, H. (1984). *Studies in Ethnomethodology*. Malden MA.
Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical statistical controversy. *Psychology, Public Policy, and Law, 2*(293-323).
Hilton, N. Z., & Simmons, J. L. (2001). The influence of actuarial risk assessment in clinical judgments and tribunal decisions about mentally disordered offenders in maximum security. *Law and Human Behavior, 25*(4), 393-408.
Krysik, J., & LeCroy, C. W. (2002). The empirical validation of an instrument to predict risk of recidivism among juvenile offenders. *Research on Social Work Practice*(12), 71-81.
Lindsey, D. (1992). Reliability of the foster care placement decision: A review. *Research on Social Work Practice, 2*(1), 65-80.
Lu, Y. E., Landsverk, J., Ellis-Macleod, E., Newton, R., Ganger, W., & Johnson, I. (2004). Race, ethnicity, and case outcomes in child protective services. *Children and Youth Services Review, 26*(5), 447-461. doi: 10.1016/j.childyouth.2004.02.002
Lyle, C. G., & Graham, E. (2000). Looks can be decieving: using a risk assessment instrument to evaluate the outcomes of child protection services. *Children and Youth Services Review, 22*, 935-949.
Mantzoukas, S. (2004). Issues of representation within qualitative inquiry. *Qualitative Health Research, 14*(7).
Munro, E. (2004). A simpler way to understand the results of risk assessments. *Children and Youth Services Review, 26*, 873-883.
NIH. (2013). Dissemination and implementation research in health (R01) Retrieved February 1st, 2015, 2015
Porter, T. (1995). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
Powell, B., Bosk, E. A., Wilen, J., Danko, C., van Scoyoc, A., & Banman, A. (2015). Evidence-based programs in "real world" settings: Finding the best fit. . In D. Daro, L. Cohn Donnely, L. Huang & B. Powell (Eds.), *Advances in Child Abuse Prevention Knowledge: The Perspective of New Leadership*. New York: Springer.
Rossi, P. H., Schuerman, J. R., & Budde, S. (1996). Understanding child maltreatment decisions and those who make them. Chicago: University of Chicago, Chapin Hall Center for Children.
Schuerman, J., Rossi, P. H., & Budde, S. (1999). Decisions on placement and family preservation. *Evaluation Review, 23*(6), 599-618.
Staller, K. M. (2006). Railroads, runaways, & researchers - Returning evidence rhetoric to its practice base. *Qualitative Inquiry, 12*(3), 503-522. doi: 10.1177/1077800406286524
Wald, M. S., & Woolverton, M. (1990). Risk assessment: The emperor's new clothes? *Child Welfare, 69*(9), 483-511.
Wilkes, J. R. (1992). Children in limbo: Working for the best outcome when children are taken into care. *Canada's Mental Health, 40*(2), 2-5.

Witkin, S., & Harrison, W. D. (2001). Whose Evidence and for What Purpose? *Social Work*, *46*(4), 293-296.
Witkin, S. L. (1996). If empirical practice is the answer, then what is the question. *Social Work Research*, *20*(2), 69-75.
Wulczyn, F., & Lery, B. (2007). Racial disparity in foster care admissions. In C. H. C. f. Children (Ed.). Chicago: University of Chicago.

Chapter 2

Same Difference: Structured Decision-Making in Practice

Introduction

The introduction of actuarial based assessments into child welfare practice represents an attempt to create a more objective framework for decision-making, reduce bias from subjective assessments, and increase the consistency of case evaluations across child protective service workers (Baird, Wagner, Healy, & Johnson, 1999; English & Pecora, 1994; Gambrill & Shlonsky, 2000; "The Structured Decision Making Model: An Evidence-based Approach to Human Services," 2008) While there is a general consensus that standardized assessments outperform clinical judgment (Baird & Wagner, 2000; Baird, et al., 1999; Dawes, 1999; Dawes, Faust, & Meehl, 1989; Grove & Meehl, 1996), limited research on their utilization in the fields of psychiatry and juvenile justice suggests that actuarial based risk assessments may be ignored in practice, restricting their value (Krysik & LeCroy, 2002; Schwalbe, 2004; Hilton & Simmons, 2001). These provocative findings raise questions about the utilization of actuarial based risk assessments in the field of child welfare. To date, there has been little research that examines how

child welfare workers integrate these decision-making aids into their every day work.

In this paper, I examine how child protective service workers understand their role as decision-makers within the context of standardized work procedures in order to build theory about the relationship between workers' conceptualization of actuarial risk assessments and their utilization. Specifically, I ask: *How do structured decision-making aids influence decision-making in the context of every day work routines? How do Child Protective Service workers make meaning of their work when much of it is guided by structured tools?* The goals of this research are to fill a gap in knowledge about how actuarial assessments are used on the ground and the factors that impede or promote their implementation.

Moving away from dichotomous understandings of standardization that frame clinical and actuarial decision-making as mutually exclusive processes, I find that workers decision-making strategies fall on a continuum. Workers perceptions of their ability to take action in a case both influences, and is influenced by, their understanding of standardized decision-making aids. Drawing from sociological theories concerned with the way that power and privilege affects a person's approach to the world, I contend that how workers make meaning of their decision-making capabilities is rooted in social factors as well as individual ones. I observe that workers who occupy social positions of higher status in terms of race, gender, and years of experience are more likely to frame their decision-making as autonomous within the context of standardized tools. The converse is also true: child welfare workers who occupy positions of lower social status in terms of race, gender, and

years of experience are more likely to frame their decision-making as less autonomous and more restricted.

Case accounts demonstrate that no matter how workers understand or construct their authority to take action, all workers maintain some level of discretion in their casework. In doing so, workers utilize a diverse set of strategies to influence case outcomes regardless of their social position or conceptualization of their professional authority. This work challenges the idea that standardized tools are able to eliminate or entirely flatten subjective processes in decision-making, and demonstrates that standardization alone cannot correct for differences in social power and standing that are a part of any social system.

In this paper, I briefly review the extant literature on decision-making in order to contextualize the challenges related to case assessment and the introduction of actuarial tools into child protective service work. I then problematize the assumption that actuarial risk assessments can be entirely neutral objects. Following this, I describe the research study and findings. This paper concludes with a discussion of the implications of this research for child welfare decision-making.

Theoretical Framing

Assessing Assessment

The difficulty of making accurate decisions in child welfare work has been well documented and represents a major concern for the improvement of the field (Dawes, et al., 1989; Gambrill, 1990; Gambrill & Shlonsky, 2000; Grove & Meehl, 1996; Lindsey, 1992; Munro, 1996). Errors in judgment can have fatal or lasting consequences for the children the system intends to protect and the families it is

charged to serve. At their most severe, flawed assessments lead to unwarranted separations between children and their caregivers or allow children to remain in unsafe situations. At an institutional level, the aggregate effects of subjective understandings of risk, maltreatment, neglect, culture, and safety as well as individual bias, lead to disproportional placement rates of children from ethnic and racial minorities in the child welfare system (Dettlaff et al., 2011; Dettlaff & Rycraft, 2008, 2010; Dworsky et al., 2010; Fluke, Yuan, Hedderson, & Curtis, 2003; Gambrill & Shlonsky, 2000).

Using case vignettes, studies have empirically demonstrated that child welfare decision-making can be surprisingly variable when workers are given an identical set of facts. Lindsey's research on the reliability of decision-making among child welfare workers found that they agreed on a case trajectory only 25% of the time (Lindsey, 1992). In separate studies, Rossi, Schuerman, and Budde (1996) and Arad-Davidzon and Benbenishty (2008) also found large levels of disagreement about placement decisions among child welfare workers evaluating the same case. There is little overlap between the criteria that child protective service workers base their decisions on and the risk factors that researchers have identified as being relevant to the risk of future harm (Davidson-Arad & Benbenishty, 2010; Dorsey, Mustillo, Farmer, & Elbogen, 2008; Osmo & Benbenishty, 2004). When workers are unsure about what to do, they are more likely to make decisions that escalate cases, interpreting the risk for future maltreatment as higher than it actually is (Davidson-Arad & Benbenishty, 2010; Dorsey, et al., 2008; Osmo & Benbenishty, 2004). Even when research exists to guide difficult assessments, workers are unlikely to

integrate this information into their decision-making (Davidson-Arad & Benbenishty, 2010).

Addressing these issues in decision-making remains a complex task. Describing the "uncertain context" within which judgments are made, Gambrill & Shlonsky (2000) summarize the "barriers to good decision-making as:

> "(1) Limited knowledge; (2) limited information processing capacities; (3) personal obstacles such as lack of perseverance, reliance on ineffective problem-solving strategies and lack of familiarity with problem-related knowledge; and (4) the task environment" (Gambrill & Shlonksy, 2000)

Child welfare cases are often messy, filled with conflicting data, and difficult to resolve. Heuristic strategies that make complex information more manageable in everyday life can lead to poor decision-making in the context of child welfare cases as they invite biases in reasoning (Gambrill & Shlonsky, 2000, 2001; Munro, 2004; Tversky & Kahneman, 1974, 2002). Case characteristics, the personal characteristics of the worker, the relationship between the worker and family, and the larger institutional context have all been shown to influence child welfare assessments, underscoring that these decisions are 'judgment' calls shaped by individual and organizational variables as well as by powerful normative ideas about 'what children need' and 'what parents should do.' (Holland, 2000, 2001; Osmo & Benbenishty, 2004; Portwood, 1998; Rose, 1999; Rose & Meezan, 1996; Walsh, Bridgstock, Farrell, Rassafiani, & Schweitzer, 2008).

The complexity of decision-making is further complicated at the institutional level by lack of definitional agreement about what constitutes maltreatment, particularly neglect. Since the mid- 1960s, researchers have noted that there is little shared understanding among child welfare workers about what necessitates social service intervention and what criteria should be used for the home removal of a child

(Barnett, Manly, & Cicchetti, 1993; Coulton, Corbin, & Su, 1995; Giovannoni & Becerra, 1970; Hutchinson, 1990; Jones, 1993; Rose & Meezan, 1996; Young, 1964; Zuvarin, 1999). Variability in decision-making should not be surprising when there is so much conceptual heterogeneity about what maltreatment "is" as well as widespread contentions about how and when to intervene when it occurs. Agency culture can make decision-making harder as workers must find a way to balance competing directives, obligations, time pressures, role requirements that call for both surveilling and supporting a family and an environment that can reward conformity in thinking over critical evaluation (DePanfilis, 1996; Gambrill & Shlonsky, 2001). The sum total of this research suggests that case decisions are more likely to be dependent on which worker a family receives and the larger institutional context in which it is processed than the objective facts of a case.

Risk Assessments in Child Welfare

In response to research documenting these inconsistencies, efforts to improve decision-making have focused on standardizing child welfare practice with the dual goals of (1) creating objective measures that will yield identical results for similar cases and (2) producing data that the child welfare system is working, evidence-based, and cost effective. The development and introduction of actuarial risk assessments represents a major intervention in child welfare decision-making. By aggregating the factors related to the child, caretaker, family, and environment that have been empirically demonstrated to be the most influential for potential maltreatment, actuarial risk assessments create a probabilistic measure for the likelihood that maltreatment will occur in the future. Transforming decision-making in child welfare from a relatively personal process to a systems-level task, structured

risk assessments standardize the criteria child protective service workers use to make determinations about their cases. In doing so, actuarial risk assessments also attempt to mitigate the racial and ethnic disparity that comprises and compromises the child welfare system. If each measure applies equally to every family no matter what their race, class, or ethnic background, then the assumption is that families will be evaluated fairly and bias will be reduced.

Actuarial risk assessments are designed not only to reduce disparities in child welfare but also improve the accuracy and consistency of child welfare decisions by ensuring that workers use the same criteria for judgments in every case; make predictions about whether child maltreatment will reoccur based on statistical models and empirical research; and to guide decisions about which families substantiated for child abuse or neglect would most benefit from services by distinguishing between families at high, moderate, and low risk for reoccurrence of maltreatment (Baird & Wagner, 2000; Baird, et al., 1999; Dawes, 1988; Dawes, et al., 1989; "The Structured Decision Making Model: An Evidence-based Approach to Human Services," 2008).

Structured risk assessments are understood to be more accurate than clinical decision-making, with over 100 studies finding that they perform better than decisions made using intuitive or clinical reasoning (Baird & Wagner, 2000; Baird, et al., 1999; Dawes, 1988; Dawes, et al., 1989; Grove & Meehl, 1996; "The Structured Decision Making Model: An Evidence-based Approach to Human Services," 2008) Dawes asserts:

> In the last 50 years or so, the question of whether a statistical or clinical approach is superior has been the subject of extensive empirical investigation; statistical vs. clinical methods of predicting important human outcomes have been compared with each other, in what might be described

> as a "contest." The results have been uniform. Even fairly simple statistical models outperform clinical judgment. (Dawes, 1993)

Over the last 15 years, these findings have led to the widespread adoption of actuarial risk assessments by child welfare agencies, the most prominent of which is the Structured Decision-Making Model (SDM), currently in use by over 30 states.

Despite the widespread adoption of actuarial risk assessments and the general consensus in the field of child welfare that these are evidence-based practices, methodological concerns about their construct validity, predictive ability, reliability, and implementation remain (Gambrill & Shlonsky, 2001; McDonald & Marks, 1991; Munro, 2004; Schwalbe, 2004; Wald & Woolverton, 1990). Response to these concerns focus on building better, more accurate, actuarial models; increasing the training of workers using them, and introducing more attentive implementation strategies.

Embedded within each of these strategies is the assumption that when designed and implemented 'right,' standardized decision-making aids can largely, if not completely, mitigate the subjectivity that impedes consistency in clinical decision-making. This paper questions the assumption that structured decision-making aids can entirely obscure the independent judgments, perspectives, and quirks that individuals bring to their decision-making in general, and their use of these decision-making tools in particular.

Research Methods

Sample

The data for the present study originates from over 70 hours of semi-structured, qualitative interviews with 35 Child Protective Service Workers (CPS),

who worked in 3 different area offices in one state. Part of a larger comparison project on how CPS workers make decisions, this paper draws from a sub-set of the total data collected in one state**.** Interviewees were identified in two ways. First, I recruited an initial sample through 3 continuing education trainings for child welfare workers held at a school of social work. At these trainings, I introduced myself, described the study and passed around a sign-up sheet for study volunteers. I then scheduled and conducted interviews with those who indicated interest. Additionally, an area director introduced my research to staff and allowed workers to sign-up for interviews at an all agency staff meeting. Following a similar procedure, I contacted volunteers and scheduled interviews with those who demonstrated interest. I also recruited participants using a snowball sample as other CPS workers heard about the study from colleagues and asked to join. Participants were offered a $25 Visa gift card to compensate them for their time.

Interviewees ranged in age from 24-62. Participants were comprised of 17 White women, 8 African-American women, and 6 White men. Workers interviewed had a variety of experience. 15 had worked in their current position less than 2 years; 8 worked in their current position between 2-5 years, and 7 had over 5 years of experience on the job. All participants were front line Child Protective Service workers tasked with either making decisions either about substantiation and severity of child maltreatment (Investigators) or, in the case of substantiated cases whether a family was complying with treatment and recommended services (Treatment Workers). For decisions about the latter group, Child Protective Service workers were responsible for either discharging or escalating cases as families wound their way through the child welfare system.

Interview Content

The interview protocol for this study was designed to capture the world-view of the participants through their common practices and actions (Charmaz, 2006). Semi- structured and open-ended, the protocol begins with a discussion about the everyday work routines of the respondents. I began the interview by asking how the participant came to be either an investigator or treatment worker and about the challenges of the job as well its pleasures. Following this initial discussion, I asked participants "to walk me through" the process in a neglect case from the time they received it to when they transferred or closed it (Weiss, 1995). This question is intended to collect a procedural accounting of casework, the 'how it should go' or the 'what I typically do.' I then asked the participant to walk me through his most recent completed child neglect case. The logic of this question is to use recent completed cases as a proxy for typical cases. These descriptions gave me a baseline for what kinds of actions or inactions were conceptualized as common in a particular office and what course these cases normally took. The next question centered on challenging cases and I asked participants to describe a case where they had a difficult time determining what actions or recommendations they should make. I then asked participants to walk me through a case where they felt confident in their decision-making or where it was easy to know what to do.

After collecting these accounts, I asked workers about the Structured Decision-Making (SDM) tools (if they had not yet mentioned them at other points in the interview). Questions about the SDM tools centered on how workers used them in their every day work routines as well as assessments of their usefulness to aid decision-making. This line of questioning ended with the worker walking me

through the Safety, Risk, or Reunification Assessments for each of the three cases they described. During this process, I was able to collect data on how the worker understood the questions on the SDM tools and how this understanding unfolded in practice. These questions also allowed me to collect data on the weight that these assessments were given in the final disposition of the case and how they influenced decision-making about legal or social service involvement. Throughout the interview, I followed up every question with probes specific to the conversation.

Common probes focused around further elucidating the logic at decision-making points and case detail. Interviews ranged from 60-180 minutes with a median interview time of 90-100 minutes. Interviews were audio taped and transcribed. I conducted interviews at neutral locations such as local libraries and a community center with offices available for sign-out. Interviews also took place in an available room at one area Child Protective Service office. All names of offices, agencies, and people are pseudonyms.

Data Analysis

Following grounded theory protocols, I began data analysis immediately after the initial interview through detailed memoing (Charmaz, 2006). Interviews were coded using an open coding strategy with individual transcripts using NVivo 10 qualitative data analysis software (Emerson, Fretz, & Shaw, 1995). I developed concepts and categories reflecting recurring themes in the data. Interviews were then analyzed at the aggregate level. After initial open coding, two Research Assistants analyzed the data deductively based on the themes and categories I initially identified. The data presented below represents larger themes found in the

data, except where noted. Quotes from the data are edited for clarity (e.g. repeated "Ums" removed), but no words have been changed or re-ordered.

Why Here

State A offers an ideal site in which to conduct this research. Currently, State A's legislature mandates the use of the Structured Decision-Making Model (SDM) an actuarial-based approach for child welfare decision-making in all casework. The Structured Decision-Making Model is one of the most prominent actuarial based assessments and is in use in over 30 states. As an early adopter of the SDM, State A's risk assessment serves as the basis for a number of other state's risk assessment instruments. In addition to the SDM, State A has also adopted a category system where Child Protective Service workers are required to rank a case based on level of severity from 1-5. The ranking system further standardizes caseworker action and discretion by specifying the specific protocols that follow each ranking. Case categories are explicitly linked to the risk assessment with scores on the RA correlating with the case category. The chart below outlines the category system and their connecting designations from the risk assessment:

State A Child Protective Services Category System

Category V	Following a field investigation, the department determines that there is no evidence of child maltreatment or is unable to locate the family. **No action taken**.
Category IV	Cases where there is no preponderance of evidence that neglect or abuse has occurred. If the risk assessment demonstrates any risk **Voluntary Community Services are Recommended.**
Category III	A preponderance of evidence that child maltreatment has occurred. The risk assessment scores as low or moderate. If the family does not voluntarily participate in services or does not progress CPS may reclassify as a Category II.

	Community Services are Needed. Moderate Risk
Category II	A preponderance of evidence that abuse or neglect has occurred. There is a high level of risk for future harm as indicated by the risk assessment. The perpetrator is listed on the Central Registry for Abuse and Neglect. **Child Protective Services are Required in Conjunction with Community Services. High Risk**
Category I	A preponderance of evidence that child maltreatment has occurred with the risk assessment scoring as high or intensive. Child Protective Services determine that a court petition is either needed or required. Perpetrator is listed on the Central Registry for Abuse and Neglect. **CPS petitions for the immediate removal of child(ren). Services Provided by Child Protective Services or Foster Care in Conjunction with Community Services. Intensive Risk**

Autonomous Workers

It's An Art Not a Science

Even within the context of the highly protocolized policy environment in State A, a subset of workers (11 out of 35) experienced themselves as having a degree of autonomy and discretion in their casework. For the purposes of shorthand, I will refer to these workers as autonomous throughout this paper. Describing their decision-making in authoritative terms and actively emphasizing their ability to make judgments about the merits of cases, autonomous workers understood themselves as able to influence the trajectory of cases and their outcomes independent of the actuarial based risk assessments that they were required to use. For this group of workers, the actuarial risk assessments were a bureaucratic tool that they had to manage but not heed. Autonomous workers' conceptualizations and critiques of the Structured Decision-Making Model (SDM),

particularly the limitations of the risk assessment, informed their view of decision-making as more of an art than a science.

Autonomous workers viewed the decision-making process, even in the context of the SDM, as highly variable. Citing differences in worker and supervisor attitudes, these workers regularly observed that case decisions and outcomes were highly associated with the personal traits, interpretations, and perspectives of their colleagues rather than measures like the SDM. Workers who observed this variability did not trust the SDM to be an equalizer across cases. Instead, they viewed the risk assessment as an administrative tool that recorded rather than mitigated diverse perspectives. Without confidence that the SDM worked effectively, workers felt that they had little reason to align their work practices around it. Decision-making about cases proceeded engagement with the SDM.

Autonomous workers reported that they first came to a decision about the case and then utilized the SDM at the end and only because that was what was required. Because the results of the risk assessment in State A mandate case trajectories, workers in State A who disagreed or disregarded its results still had to negotiate a strategy for the SDM that would preserve their intentions for a case. These workers described calibrating their answers on the risk assessment to reflect decisions they had already made or relied on familiarity with the risk assessment itself to interpret questions in ways that would either raise or lower the final risk level. Interpreting questions with an eye toward the final risk level, in turn, affected the category at which a case is substantiated. Compared to workers with a more restricted sense of autonomy, workers who understood themselves to possess discretion in cases, viewed themselves as having multiple points in the investigative

or treatment process during which they could intervene to shape the way a case unfolded including the actual filling out of the risk assessment itself.

Workers who both doubted the reliability of the SDM and also felt free to make case judgments independent of the information contained within it, generally were in a position of privilege in terms of either race, gender, years of experience or a combination of the three relative to other workers who shared their critiques of the SDM but who nonetheless understood themselves to be restricted in their ability to influence case outcomes.

An Extremely Poor Tool

Autonomous workers repeatedly described the risk assessment as an inaccurate and/or ineffective tool that did not play a part in their decision-making. Rob explains how little he relies on the risk assessment during his routine casework:

> E: Can you tell me about the risk assessment and how the risk assessment works to influence or sort of how it impacts decision-making?
>
> R: Well I hate to start by saying it's an extremely poor tool. It's an extremely poor tool that...(looking at risk assessment)...Yeah I could do these in my sleep. We see those forms so many times. You breeze through them. You're not paying attention...You know the other day, it was like probably the first time in a long time, I had a case and I wanted to know; it's either a Category 2 or a Category 3. My boss was like, well do the risk assessment as completely and thoroughly as you can and see what it spits out. You know it came back as a 2. So, that was like the one time I relied on the risk assessment to give me kind of what CPS thinks is a Category 3 or Category 2. (White, Male)

For Rob, the risk assessment rarely impacts his decision-making as evidenced by the fact that he can only remember one time that he used the RA to determine the category for a case. Having seen the RA "a thousand times," Rob can accurately gauge how to fill it out so that it reflects decisions he believes are justified. Instead of being a powerful factor in case outcomes, the risk assessment is relegated to a

bureaucratic tool, which requires little thought and functions without much merit. Rob goes on to detail how he understands decisions to be made:

> R: I'll give you the off the cuff. When I do a case and I put together a packet, we have a case conference. I know sometimes the workers will read the 154. Nobody reads these [the risk assessment]. Nobody reads these. They're in the file. It's just a piece of paper that the state can say 'hey that we did this.' The workers will read the 154 and they will take the information we give them. So whatever they're doing, they're basing it off of what they read in my 154 or whatever I tell them in that meeting when we're transferring the case. Me and my supervisor and them and their supervisor. So, these have no; I mean yeah the risk may give us some guide about whether it's going to be a 2 or a 3.
>
> E: Do you find that helpful ever?
>
> R: What, the risk?
>
> E: Yeah.
>
> R: No. (White, Male)

Not only does the risk assessment provide little useful guidance about case decision-making but also, according to Rob, it is not used reliably by others to interpret information about the case. Instead case narratives, case conferences, and supervisory discussion are the methods through which a case trajectory begins to be framed and a course of action takes shape.

Other workers share Rob's experience of the risk assessment as an inaccurate tool and in doing so, come to similar conclusions about whether it should factor into decision-making. Skeptical of the RA's accuracy, Jennifer does not refer to it when formulating her decisions about her cases:

> J: ...because the risk. I mean yeah, sure you have to have a risk and a safety assessment but I don't know how really accurate they are. They're really general questions and they're really, um, they're just, I don't know, they almost like seem like they're too generic... I'm not a big fan of them.
>
> E: So when do you do them in the course of your investigation?
>
> J: You're supposed to...Well you're supposed to do the safety assessment once you commence your referral and figure out the safety to the child, (short laugh) but someone like me, I do them at

the end once you know, so once I'm ready to dispo my case, (short laugh)

E: So you do them at the end?

J: Yeah I do. I'm not going to lie about that cause I can't...

E: No, and you're not alone either.

J: Yeah I'm like oh I got to dispo my case, I'm like doing the safety, I do the risk, well FANS [Family Assessment of Needs and Strengths] and CANS [Child Assessment of Needs and Strengths] is only when you have to do an open case, but yeah I generally...

E: So do they impact or influence your decision making in any way?

S: No. (White Female)

Like Rob, Jennifer sees the risk assessment as a poor proxy for clinical decision-making. Lacking nuance or complexity the risk assessment and other parts of the SDM do not satisfy Jennifer's criteria for the kind of information she needs to make informed decisions about her cases. William shares Jennifer and Rob's assessments noting that "I would be very, very hard pressed to prove if we want to talk stats for a second... it [the risk assessment] is not valid and it is not reliable." Without confidence in the accuracy of the risk assessment or confidence in its ability to meaningfully inform cases, workers like Jennifer, Rob, and William feel justified in using other methods for decision-making than the protocols or policies the Department of Health and Human Services requires.

Objectivity is a Subjective Experience

What makes the risk assessment a poor tool in the eyes of some workers? State A workers who understand themselves to maintain professional authority and discretion share the same set of critiques about the SDM namely that: The risk assessment (1) lacks nuance; (2); is unable to effectively distinguish between moderate and high risk cases; (3) penalizes families for past history rather than current issues; (4) functions primarily as a bureaucratic tool for management; and (5)

and is easily open to multiple interpretations which undermine its claims of objectivity. It is this final factor that workers who maintain a sense of professional authority and discretion focused on the most when explaining their resistance to the SDM and their approach to decision-making. Rob sees the risk assessment as just another place in case work where individual interpretation guides decision-making:

> You know I tell people this all the time. If you had one CPS case and it has a borderline complaint, you know could be, couldn't be. And you gave it to 50 different workers, all throughout the state with different supervisors, you'll probably get 20 Cat 4s, 20 Cat 3s, a couple Cat 2s and then maybe one or two guys will give it a Cat 5. So it's like, so you say how do you come to a decision? You know it depends on what worker you got, what supervisor that person is under because the supervisor sets the tone for what kind of mindset the worker is going to have. You know some supervisors are more aggressive and less trusting and you know, more aggressive when it comes to I guess punishing a parent or what have you. (White Male)

Concurring with this assessment, William enumerates how the questions on the risk assessment themselves invite multiple readings, which leads to disparate risk levels for the same case:

> *Primary caretaker currently has a substance abuse problem.* One could read that one of two ways. There are workers that will say, "Okay, they're in treatment." They still have a substance abuse problem because they're in treatment. So that's a yes, which is immediately going to make the risk higher. But they voluntarily went in treatment and they've completed treatment; so is it still a problem? I mean and you know, okay part of that could be a little semantical, but if you really read that; I mean the way I read it would be; yeah, they did have a problem, but they voluntarily have participated and the worker says they're doing absolutely fantastic and there's no problem with that at all. So, I would answer that "no" but others would answer it "yes" and that's immediately going to skew things.
>
> Oh this one's beautiful; *primary caretaker views incident less seriously than the department.* That is so wrought with lack of objectivity; it's not even funny. You know and then that's where the worker has to really deal with their own personal feelings and their own biases. (White, Male)

Rob and William both see the risk assessment as capturing the diversity of workers opinions rather than as a tool for preventing variability. Consciously or unconsciously workers own perspective on norms related to parenting, child care,

substance abuse, substance abuse recovery, exposure to domestic violence, parental insight, and a family's cooperation with the department influence how a worker will approach the interpretation of questions on the risk assessment. Diverse understandings of what constitutes risk come to light as each worker deciphers each individual question. As a result, cases with the same factual evidence can follow very different paths based on how a worker interprets the questions on the risk assessment.

Workers who understand the risk assessment to encode subjectivity rather than objective levels of risk then felt free to approach the risk assessment as a place to document rather than uncover decision-making. Workers described how both they themselves and their colleagues adjusted answers to alter the final score of the risk assessment. Because these workers understood the RA to capture variability, they felt they were not doing anything differently than their colleagues when they used flexible and distinct interpretations of questions to guide case outcomes. Disputing the idea that an actuarial approach successfully eliminates subjective responses, Jackie stated: "I get that this is numerically based. But, clearly you can manipulate it." Matt describes both the rationale for altering the risk assessment along with a description of how such manipulation takes place:

> E: How do you use those tools [the SDM] in your, when you're working on a case?
>
> M: They're just blockades, I mean I just get 'em done and it doesn't even mean anything to me, because I know what the risk is. I'm not going to finish the thing and look and go, oh it's moderate. I mean, I know, and frankly just about everything comes out as moderate anyway unless you got some severe stuff happening. And in that case you already know it. So I don't use those for guidance at all. And surprisingly, when I do those, and I don't even look to see what level category it comes out at. I just start working on the disposition the first sentence is, this is a category 3 case. And guess what, 99.9% of the time it's exactly what I thought it was gonna be.

E: And what happens that .1% of the time when it's not?

M: Well then I've got to go back and fix it. (Laughter)

E: Oh okay, and then you've got to go back and fix it.

M: Yeah, either I will change the risk assessment 'cause I will go through it again and say, you know what that might have been a problem, but it's not that big of a problem. Or I will change category 4 to category 3 or category 3 to a 2. (White, Male)

Matt performs his professional acumen by demonstrating how often his understanding of risk aligns with the understanding of risk encoded within the risk assessment. The risk assessment is a blockade precisely because he believes he assesses risk as well, if not better, than the SDM. However, when Matt's understanding of risk conflicts with that of the risk assessment it is not a cue that he got the information wrong but an error within the risk assessment itself. Matt "finagles" the results because of his confidence in his professional authority. William too notes that the risk assessment often needs refinement and sees adjustment of the risk assessment as an issue not only of professional expertise, but also of morality:

E: How often would you say that you have a case where your assessment of what the risk level is or the category should be is in conflict with what the risk assessment has scored?

W: I would say right now; actually in my experience, literally 60 percent of the time.

E: So more than half you're doing something different than what the...

W: These [the RA] are making decisions on people's lives and I want to make sure that it's going to reflect what I know and what I've observed and what I've assessed. (White, Male)

Because William understands the risk assessment to be inaccurate, it is important that he amend his interpretations on the risk assessment so that its outcome reflects his assessment of the risk level for the case. As William points out, the stakes for families are too high to proceed with a finding that ignores his clinical experience. For William, being a professional means using the SDM interpretively, privileging

his years of experience as a social worker over what he sees as an ineffective form. Doing so is not unethical, but rather the only ethical course of action.

Workers consistently cited the serious consequences of an inaccurate (or even an accurate) risk level for families as the primary reason that they revisited and reformulated the risk assessment. Jackie explained that the placement of a parent or caregiver on Central Registry was one area that she and other workers took into consideration in how they scored each case:

> J: I think that with cases where you know it's gonna be substantiated, I think that you do think about Central Registry and I think that's probably one of the most manipulated areas and that you can manipulate it. Because while it's based on the risk, and what you put in for the risk, what you don't put in for the risk or just the way different people view, you know if, one of the questions is something like; is there domestic violence I could say in their current relationship, no. But, mom was beat by her former husband twenty times and the police have been out there, so I mean technically I could mark yes. So, I mean some of those things can be manipulated; they really can.
>
> E: Is it something that you think about when you have a case that you know could approach a level 2 or a level 1?
>
> J: Absolutely (White Female).

Jackie understands manipulation to take place because the risk assessment itself is so open to subjective understandings of events and the circumstances of a family. This experience of the SDM as well as an awareness of its consequences facilitates workers critical attention to the relationship between the facts of the case, the score of the risk level, and how the worker believes she should proceed. Like William, Jackie believes that adjusting the risk assessment to reflect her understanding of the case is an important professional act that she preforms in the course of her everyday work. Being able to use her assessment skills to decode the risk assessment is a part of the work that Jackie has been hired to do.

Ultimately, the workers that act to maintain a high level of professional autonomy in their work despite policies explicitly designed to constrict it, have constructed their role as professionals in terms of their ability to make decisions. As William put it: "I mean in all my years, my conclusion has always been that this field requires a ton of experience, learning from your mistakes, and it's much more of an art than it is a science; period." These workers reject quality improvement measures like the SDM because they do not believe that quality decision-making can be done without privileging clinical reasoning.

Workers' assessments of the limitations of the SDM mirror concerns about methodological weaknesses of actuarial based approaches. Static variables on risk assessments such as number of children, age of children, etc. are often rated the same by different workers evaluating the same case but items that require more interpretation such as determining whether a parent is cooperating with the worker or the complaint of maltreatment is being taken seriously are often rated variably depending on the analysis of the person filling out the risk assessment. Beliefs that the risk assessment is not valid or reliable reflect concerns about the external validity and predictive power of actuarial based assessments in use in the field.

Suggestions to address these issues have focused around digitizing the decision-making aids, increasing training related to the meaning of each question, and providing examples that highlight the kind of clinical reasoning that the question attempts to encode (Gambrill & Shlonsky, 2000; Shlonsky & Gambrill, 2001). In State A, all of these suggestions have been enacted: Risk assessment are completed and calculated on the computer using the SWSS system; workers have access to definitions and examples of case thinking for each of the items on the risk

assessment; and workers are trained extensively in use of the risk assessment. In practice, attendance to implementation issues has done little to mitigate the challenges of standardizing diverse thinking and experiences among child welfare workers who argue that the flaws they see in the risk assessment justify ignoring it when the results do not match their understanding of the case.

Implementation issues are often framed in the literature on evidence-based practice in general, and risk assessment utilization in particular, as related to workers not understanding the importance of using evidence-based methods or valuing practice-wisdom and intuitive reasoning over scientific methods. It is important to note that, here, issues of implementation are, in part, related to concerns about the validity and reliability of risk assessments and that some of these concerns have also been raised by experts in the field. Workers' critiques of the risk assessment and stated reasons for subverting its intentions invite a more nuanced conceptualization of implementation issues than simply improving training can address. Without direct attention to workers concerns about the limitations of the risk assessment, it is unlikely that any intervention that only addresses the process for using the SDM will succeed.

Restricted Decision-Making

The majority of workers in State A constructed their decision-making in restricted terms (24 out of 35). These workers felt that they were not in control of a number of choices that were directly tied to the score a family received on the risk assessment (RA) such as determining whether to place a perpetrator of child maltreatment on the Central Registry for Abuse and Neglect, the severity of a case,

and the intensity of treatment needs. Workers who experienced their decision-making to be restricted fell into two groups: those who also conceptualized the RA as an ineffective tool but who felt they had no choice but to comply with the policy and those who believed the RA was an important aid to decision-making. Compared to workers who conceptualized themselves as having autonomy over their casework, workers with a restricted view of decision-making (whether positive or negative) tended to be in minority positions with respect to race, gender, and years of experience.

Workers who experienced the SDM negatively and who also conceptualized their decision-making to be restricted (11 out of 24), felt that this lack of control greatly diminished their professional identity. I will refer to these workers as negatively restricted. For these workers, limitations imposed by the SDM meant that their cases could unfold along trajectories with which they disagreed and which almost always undermined their professional authority. Among negatively restricted workers, there was a fatalistic sense that 'nothing could be done' to correct outcomes to which they objected.

For workers who held a positive view of the Structured Decision-Making Model (13 out of 24), there was little conflict about restrictions in decision-making .I will refer to these workers as positively restricted. These workers expressed relief that there was a scientific process for making case judgments (the risk assessment) that relieved them from the burden of responsibility, making mistakes, and making choices that could alter the construction of families. Characterizing their restricted authority as being necessary for carefully carrying out their work and reducing bias, these workers accepted the risk assessment unquestioningly. As they did so,

positively restricted workers redefined the meaning of being a professional from a person able to make autonomous decisions and take action to a person that is in charge of following established rules and executing policy properly. Utilizing a standardized protocol for decision-making allowed these workers to manage the uncertainty and risk inherent in their choices without being overwhelmed by the responsibility these choices entailed.

However, all workers with a restricted sense of discretion (whether positive or negative) were not entirely without autonomy. Finding places to exercise discretion in their casework outside of the SDM and the risk assessment, restricted workers preserved other spaces in the life of a case to guide its trajectory. These workers described using the 30-day investigative period as the site where they were able to influence case outcomes. During this period, workers evaluate whether there is a preponderance of evidence necessary to substantiate a charge of child maltreatment. Decision-making related to case substantiation gave restricted workers a place to exercise professional judgment and offered opportunities to take action outside the scope of the SDM. Interventions occurred in two forms either (1) helping families that they did not want to see enter the child welfare system correct issues that otherwise left unattended would mean a substantiation; or (2) by separating a child from their family through informal safety plans before the investigative period concluded. In this way, discretion was not entirely restricted; it just shifted form as workers exercised their authority in an area of case decision-making not guided by standardized tools

I Don't Make Decisions, The Computer Does

Workers, who experienced their own decision-making power as restricted, consistently constructed case decisions guided by the risk assessment to be entirely outside their own power and dictated by another entity. Asked to explain how they make distinctions between categories when evaluating cases, workers who understood their discretion to be restricted described being removed from the process. Maddie's explanation of decision-making typifies how workers constructed decisions as outside of their control:

> E: How do you understand the difference between Category 3 cases and Category 2 cases?
>
> S: You know, and the crazy thing is too, we don't even make the determination, the computer makes it for us so it's like... (White, Female)

The computer here is the digitized version of the risk assessment, which is calculates scores based on the answers a worker provides to its questions. Maddie sees herself as separated from the decision-making process without any sense of where the final determination actually comes from. Decision-making is not attributable to Maddie or another person (e.g. a supervisor or administrator) nor is it related to the information provided by the risk assessment. Instead, Maddie conceives of case decisions tied to the risk assessment as emerging from a machine, lending decision-making a mechanistic quality rather than a processual or agentic one. In this formulation, the worker is a bystander rather than a participatory actor. Dorothy, Casey, and Tia also describe decision-making in these terms:

Dorothy:

> E: And how do you assess that it's like a Cat 3?
>
> D: When you go to your dispo it'll ask you; you know is there a preponderance of evidence? Usually when you do that one, if you

click yes, it'll do it as a Cat 1 or a Cat 2 depending on the risk level; the risk assessment. (African-American, female)

Casey:

C: Yeah. We have this computer program that's in our SWSS program that somebody created and it asks questions that they feel are risk and safety assessments; and we just check boxes yes or no, and we answer why we answer yes or no; then out pops a number that says what level of risk they have and that is what it's based on. (White, Female, My Emphasis)

Tia:

E: Right, got it. So, it sounds like for all other cases to determine the severity level, you really use the forms [the SDM tools].

T: Yeah, the forms.

E: Like the forms will guide you and sort of will help you figure out what...

T: Well, we really don't have a choice; it does it automatically for us.

So once we press disposition, its done, it pops up with this thing; well this is a high risk case and it's a Category 2 and the person will be put on Central Registry (African-American, Female, My Emphasis)

Each of these women conceptualizes her role in relatively big investigative decisions as deeply restricted. The computer (or the risk assessment) doesn't take into account what the worker has heard, felt, experienced, and thought about the case. Instead the worker feeds information related to the risk assessment into a computer and waits for a final result to "pop" out, telling them how to proceed without offering an opportunity to present their own understanding of the case or the ability to change course if it's a decision with which they disagree. While Casey presents a slightly more nuanced understanding of the way the case categorization is determined by joining her Yes or No answers on the risk assessment to the outcome, she still see herself as outside of the final process.

The computation of the risk assessment leads workers to feel as if they don't have a say in the trajectory of their cases and to experience their decision-making

capacities as inherently limited. Workers, who construct case decisions emerging from the risk assessment to be automatic, de-emphasize the connection between how they interpret the questions presented on the RA and the final result. In this way, decision-making is decontextualized and the choices that a worker makes throughout the investigative process become obscured or are not perceived to be choices at all. As a result, decision-making becomes conceptualized as highly restricted.

Negatively Restricted: Undermining Professional Identity

For negatively restricted workers, the restrictions on decision-making encoded within the risk assessment are understood to be an impingement on their ability to carry out their jobs effectively. Casey describes how hard it is for her to move forward with case decisions that she opposes:

> C: I guess to say how many times it happens where it changes what I would do because with the computer program; I would say probably it only happens like once every other month for a case. But when it does, it's just, I don't know; it's very difficult for me to sign my name on it when I don't agree with it. (White, Female)

Even though conflict with the SDM does not often emerge for Casey, when it does, she does not resolve it by "finagling" the result as some of her coworkers with a greater sense of autonomy and discretion have done. Instead, the decision informed by the risk assessment stands and the trajectory of the case becomes difficult to accept, challenging Casey's sense of professional autonomy.

Negatively restricted workers feel that there is very little that they can do to influence or alter the outcome because of the constraints imposed upon them by policy. Lila recounts her actions on a case where she believed that the caregivers should be substantiated for neglect but felt that the category to which they were

assigned (based on the score of the risk assessment) was more severe than the evidence warranted. Concluding that the family did not need as high intensity an intervention as their case categorization dictated, Lila frames her decision-making about how to proceed in terms of a moral dilemma with a clear answer:

> L: ...so but it came out like based on these [the risk assessment] and the Safety Assessment, it was...
>
> E: ...that's very high... I'm just looking [at the score].
>
> L: I couldn't... Listen, I won't lie you know what I mean or fudge it. Like based on my interpretation of that [the facts of the case] and I went through that thing [the risk assessment] like three times, but there is nothing that I can do to avoid the truth. (Non-White, Female)

Despite her efforts to "redo" the risk assessment, Lila is left with an outcome with which she ultimately disagrees but which she feels cannot be disputed. Lila does not feel free to privilege her own clinical judgment over the conclusion of the risk assessment as some of her colleagues do, so instead she must privilege procedure. Lila's commitment to following procedure means making sure that the truth or the facts of the case are accurately represented, even if her interpretation of the meaning of those facts for the risk of future harm to the children is qualitatively different than the conclusion the SDM draws. Angela offers a similar perspective:

> Sometimes you just feel bad. That's all...you just feel bad and you try to look...like you try to go over the risk assessment again and see maybe if you did something wrong if you clicked on the wrong button. And it's like I don't...'cause it was a case that I opened for a dad and the mom is the one who had this previous history and you know he just married this woman and he did give the kid a whipping and they got a bruise; or something like that. And I felt like, "Oh well he'll be a Category 3 since it's just him. But even though I did a separate risk assessment for him, they still made him a Category 2 because the mom had all this previous history. And so he...[ended up on the Central Registry](African-American, Female)

Like Lila, Angela may disagree with the outcome of a risk assessment and its attendant consequences but adherence to procedure requires that she go along with the outcome. Even though Angela believes that the father in the case does not

represent a high risk of being a repeat offender of maltreatment, she feels that there is nothing she can do to change the course of the case without violating policy. Unlike their colleagues with a greater sense of autonomy and discretion, Lila and Angela do not adjust their answers to questions on the risk assessment to change the final risk score and trajectory of the case. Instead, their strategy is to revisit the risk assessment again to make sure that they have not missed any information that could alter the final outcome. `

Workers who understand their decision-making capacities to be restricted and who feel undermined as a result, often shared critiques of the ineffectiveness of the SDM with their colleagues who felt entitled to change the results of RA. Jade explains that her sense of professionalism is challenged by a decision-making aid that she feels lacks nuance and a clear ability to accurately distinguish between moderate and high risk cases.

> E: I just wanted to note for the recorder that you sort of rolled your eyes a little bit.
>
> J: I think the system is really stupid and I don't understand who created it and there's no discretion whatsoever involved [in decreasing a category] downwards. You can increase it on your own if you think that. But for instance the risk assessment. The things that make you high risk, some of them have nothing to do with you, like you can be a victim of circumstance and things are completely out of your control.
>
> E: Can you give me...
>
> J: And you have a high risk, for instance if you have, I guess having kids. But just for argument's sake: If you have more than 3 children you get a point. If you were a victim of child abuse, you get a point. If you are a victim of domestic violence, you get a point. If your child has a mental health issue, you get a point. If your child has behavior issues, you get a point. Just things that it happened to you or your child has or there's things that can make you have... I mean obviously there's other things such as if you don't think that this is serious and you deny the seriousness of the case or if you refuse to cooperate with services or if you yourself have a criminal history. Things that you have done that affect it.

> But I've had cases in the past where, only the boxes that the victim of circumstance were checked or the kids had mental problems; and they ended up being a high risk case and I would like to just you know offer services and deny the case, but instead I have to make it a CAT 2 and put them on Central Registry.
>
> And this woman in this particular case had a job with children so she was at risk of losing her job because somebody created a program that said, because she was raped as a child, had a bad husband that beat her, and because her child had serious mental health issues and she has four kids, that she's a risk to other children. (African-American, Female)

Jade understands the SDM risk assessment may deem a person at a higher risk for future child maltreatment based on factors that may or may not influence the current risk level, such as demographic information, past history with violent relationships and whether a child has any developmental delays or mental health issues. Many workers identify these factors as being outside a person's control and not always, if ever, relevant to the need for services and/or the likelihood that a person will abuse or neglect a child in the future. When a number of items that a worker believes to be unrelated to a caregiver's behavior are endorsed on the risk assessment, the results are interpreted as being artificially inflated. For Jade, the lack of connection she sees between certain items on the risk assessment and the actual severity of a case means that she does not trust the risk assessment itself to always be correct or to accurately distinguish between moderate, high risk and intensive cases. The gap between how Jade would proceed with the case (substantiate but offer voluntary services) and the trajectory the case follows (substantiate, require intensive services, and impose a punitive consequence) reinforces a feeling of lack of professional discretion that is ultimately experienced in deeply negative terms.

In the case that Jade describes, the mother looses her job as a result of Jade clearly reporting the information she has been given during her investigation, making Jade an unwilling participant in the outcome. Like Casey, there is nothing that Jade feels she can do to change a result that she sees as unfair. Angela similarly both critiques the external validity of the risk assessment and explains how powerless she feels when the risk level is high based on factors that she believes are unrelated to either current or potential child maltreatment:

> A: Well sometimes with the outcome of the risk assessment I don't agree with; like cause you know when it comes to high and it's open, it's gonna be a Category 2. And sometimes I feel like it shouldn't be 'cause sometimes it's just for having four or more kids in the home... then that throws it over the edge to make them be a Category 2 and puts them on the unfit list; even when it's like a first-time offense.
>
> E: What do you do in those situations? 'Cause it sounds like...
>
> A: There's nothing you can do. (African-American, Female)

Like Jade and Casey, Angela has no way to resolve the conflict between what the risk assessment constructs as true and the truth that she uncovers during her investigation. Angela possesses a fatalistic assessment of how this discrepancy impacts her cases, understanding that her restricted discretion prevents her from influencing the trajectory of cases where the risk assessment scores higher than she believes is justified.

This fatalism extends to situations where workers with a restricted sense of discretion see a potential to change the policy to allow other outcomes in similar circumstances. Lila describes a case where the Risk Level is high and as a result she is required to file a petition with the court to remove the child from the home. While Lila believes that the mother should not be caring for the child, she also believes

that court involvement could be avoided since there are many other adults to care for the child at home:

> L: Yes, yeah like getting court involved and removing the child. I really didn't like there was no, how do I say it, with autonomy?
>
> E: Go for it.
>
> L: Well, no because the issue is, there was potential for harm or risk, but he was surrounded by a plethora of adults, honestly mom was not really left alone. Great grandma was there, grandma was there, dad was there, aunts were there, her brother stepped in, like mom's brother so uncle came along. Like there was enough family support that we could, you know continue to try safety planning before involving that [the court] and giving dad an opportunity to get something in order... (Non-white, Female)

Lila's professional assessment of the mother's ability to care for her child is aligned with that of the risk assessment. However, the risk assessment is unable to accommodate the other information that Lila possesses about the family and the child's risk for future maltreatment while in the care of his extended family in a home his mother also shares. Lila's sense of her lack of autonomy does not permit a solution that she feels both is reasonable and minimally disruptive.

The limited opportunity for nuance available on the risk assessment combined with a worker's inability to make decisions that are reflective of their assessments contributes to this set of workers negative experience of the SDM. Not only do negatively restricted workers feel that their authority is limited but they ultimately do not have faith that the risk assessment accurately captures the information that is truly necessary to make sound child welfare decisions. Casey highlights the implicit conflation built into how child welfare decision-making in State A is structured:

> C: ...we lose discretion, I feel like, by this assessment program. I've explained... you know I've had this conversation with my supervisor before and they often explain to me how it's suppose to figure out the child abuse risk potential and I understand that idea, that this is

> the potential for being a child abuser. But I feel like our overrides should be able to able to go both ways. I feel like, cause I know people, you know in my own personal life that meet a lot of these things, and maybe they'll be a high intensive risk, but you know, just because you're a victim of domestic violence or because your children have issues or because you have, you know four children or mental injury, you know things like that; you don't always have to fall into the statistic of what's likely to happen. You know you can overcome those odds and be a good person and be a good parent. So I just wish that we had the ability to override both ways, that's all. (White, Female, My Emphasis)

By tying the risk assessment to determinations about the severity of a case (the case categories), risk level becomes a proxy for the seriousness of the maltreatment that has already occurred (and which was the impetus for the investigation itself). There are multiple problems inherent in such a conflation: neighborhoods where maltreatment investigations are more likely to occur are also more likely to possess the demographic features that would inflate the formal risk assessment (e.g. more than 3 children in the home, previous history of domestic violence, previous history of substance abuse, higher incidence of children with mental health issues). This means that in specific populations or neighborhoods, risk levels are likely to be high and in turn, the consequences are more likely to be severe, no matter what the initial complaint for maltreatment was or what a worker believes the potential for future harm is likely to be. The risk assessment itself is not predictive of future behavior, and as Casey points out not all people with specific risk factors, even those who have been substantiated for an instance of child abuse or neglect, are necessarily going to go on to maltreat their children again.

The sense from workers with a negative understanding of their restricted autonomy is that the risk assessment itself is not a fair tool either because it lacks nuance, is not accurate, or too closely ties the past to the future. Without the belief that the risk assessment effectively measures the risk for future harm, accurately

defines the severity of the current case, and then properly metes out punishment or corrective actions in response; workers cannot experience their restricted decision-making in anything but negative terms. In this way, how workers understand the SDM to function on the ground influences how a worker responds to structured decision-making, which in turn, affects how she makes meaning of her decisions.

Like their autonomous colleagues negatively restricted workers articulated concerns about the predictive validity of the risk assessment that are shared by researchers. Gambrill and Shlonsky (2000) note that in the development of actuarial risk assessments the factors intended to be predictive of the likelihood for future maltreatment may not be reliable or valid or particularly sensitive. A validation study of a risk assessment used by Minnesota raised questions about the ability of the assessment to clearly distinguish between low and high risk cases (Loman & Siegel, 2004). Workers insight into the limitations of the risk assessment informs the larger theoretical debates about their validity, sensitivity, and reliability.

By design; a worker in State A is allowed to initiate an override to raise a risk level they deem to be too low but is not permitted to lower a risk level they believe to be too high or unreflective of the actual risk. Because case categorizations are tied to the risk level, mitigations for the risk of future harm do not have a place to be captured in the formal protective services assessment process.

Positively Restricted: Re-Defining the Meaning of Being a Professional:

Some workers with a limited sense of discretion do not feel that restrictions on their professional judgment are inherently problematic (13 out of 24). I refer to these workers as positively restricted. Positively restricted workers frame these

limitations to be both important and necessary to carry out the work they do well. LaTonya explains why such restrictions are valuable:

> We don't choose the categories so, and that's the good thing because I know some people think that "Oh, she didn't like me, so she did this." And, actually its computer generated. Like, so we don't set it. It's based off of their risks and the allegation so it's not based off of our own personal feelings um, so as far as if we set it at a higher level or not, it's really based off of the risk to the children and the family. (African-American, Female)

LaTonya echoes the language of evidence-based practice to make a case for the SDM being necessary to eliminate potential sources of subjectivity and bias that stem from "personal feelings." By having the responsibility for certain case decisions (such as the selection of the case category) taken out of her hands, LaTonya feels more credible in her position. LaTonya can assure clients and herself that her decisions are professional rather than personal, that they are based on objective evidence about risk for abuse and neglect rather than her own interpretation of how a family is doing. Keisha's description of how she completes the risk assessment makes a similar point:

> I can have parents that are totally assholes and didn't even want to cooperate with me; I still take that same 30 minutes to 2 hours sitting there, doing that risk assessment because they're just mad because I'm out there. They just mad because I'm in their life. But does that mean that you know they doing this, really, to this child? I don't base it on none of that. (African-American, Female)

Keisha underscores her commitment to fairness and objectivity by explaining how she approaches the risk assessment with challenging caregivers. Keisha trusts the risk assessment to offer an accurate view of the family's risk level and service needs, so she makes sure to fill it out carefully no matter how the family has treated her during the course of the investigation. In doing so, Keisha defines her professionalism in terms of how precisely she handles the SDM. Keisha understands

that the risk assessment carries with it an enormous power in the investigative process for determining the trajectories of cases. She told me that

> "My risk assessment, 9 times out of 10, is gonna be the most accurate out of my whole report... I don't play with that. That's serious. That's someone's life. That can mess up somebody for the rest of their life." Because Keisha views the risk assessment positively, she organizes her case decision-making around it, which in turn, is aligned with agency policy.

Positively restricted workers often frame their professional role in terms of being able to execute policy properly rather than impose their own judgments and opinions on a family. Explaining how she approaches her investigation, Gabrielle frames her professionalism in terms of her ability to evaluate a case without imposing her own clinical judgment on the situation at hand:

> If there's a bruise we don't say that that bruise is inappropriate, we say "go to the doctor and get evaluated" and that doctor makes a decision. So I never go in and say what I think. I say I'm gonna take a picture of this home, look at it with my supervisor and we'll determine... Basically I talk about policy. There's a policy for everything, so I know the policy, I have the policy and I basically read them the policy. This is not [Gabrielle] this is the policy that states that if we find this, then this has to be taken care of. If I feel like the house is not up to my standards I don't judge it, I go by policy. If policy says you know, if there's feces on the floor or wires hanging or no water, then that's what I deal with. I could care less if it's clothes all over the floor. Is there a path where you can walk? That's what policy's says, so that's my position on it. (African-American, Female).

Gabrielle not only understands her professional role to be tied to how well she is able to know policy and execute it, she actively frames herself as lacking in discretion and autonomy when it comes to her cases. For Gabrielle, none of the decisions she makes come organically from her ideas, opinions or evaluation of the circumstances of the case. Others verify the evidence (a doctor, her supervisor) in accordance with *their* professional judgment. By making others responsible for the decision, Gabrielle redefines her professional role in terms of being an arbiter of fairness. Acting only when the policy requires and not before, Gabrielle sees her

work as important in that she is able to eliminate a subjective assessment from her investigation. In this way, Gabrielle redefines professionalism for an investigator in terms of the absence of autonomy and discretion.

At the heart of each of these constructions is a formulation of professionalism that is in conscious opposition to critiques of caseworkers as biased, subjective, and out of touch with the evidence that forms the heart of any case decision-making. These workers instinctually or purposefully move away from descriptions of work practices that might open them to censure for 'not knowing enough' or for being too emotional in their casework. The lack of discretion in the work environment becomes an asset rather than a liability as workers perform their authority by having a clear command of the rules. Defining a new space within which to engage their work these caseworkers refashion not just the meaning of their investigative process in the context of the SDM but also their work practices. Shannon notes that she organizes her interviews to make sure that she captures all the information on the risk assessments:

> With the risk assessments I do, not all the time, but on most of my cases I take the risk assessment and I print out this and I take it out with me and I just go over; like I ask parents questions like deemed towards this.

Shannon, like Gabrielle makes sure that her work practice aligns perfectly with work policy, letting the policy guide how she approaches her investigation, what information is the most important to obtain, and how that information is processed. Shannon's attention to detail in this area communicates that her performance as a competent professional is tied to her incorporation of work policy into her own work practices.

Other Avenues for Professional Authority

The workers discussed above all experienced themselves to have limited discretion in their work practices and in relationship to case outcomes. While it would be easy to draw the conclusion that these workers had no ability to work in ways that reflected their own beliefs and judgments, I find that the picture is much more complex than simply 'the SDM restricts workers ability to influence their cases.' While the workers above all perceived their options for response to the SDM as inflexible, workers in this sample, whether they constructed their restricted authority in positive or negative terms, all were ultimately able to maintain some measure of professional autonomy when it came to their cases. Discretion did not completely disappear but instead found other permissible avenues for expression.

According to DHS policy, the outcome of the risk assessment does not determine whether an accusation of maltreatment is substantiated and a case opened, it only establishes the case category. The decision to substantiate a case is based on whether a preponderance of evidence exists, a determination that is largely left to workers and supervisors to conclude. Workers who felt they had few opportunities to influence the trajectory of a case once it was substantiated (read: those with a restricted sense of authority) used the 30-day investigative period as the site to exercise their own judgment and offer the kind of professional intervention they believed were most likely to alter the course of a case. Angela explains that she doesn't have any control over the risk assessment and the Case Categorization but that she does have control over whether a case opens:

> E: So then how do you make a decision?
>
> A: But if we do substantiate it, then it'll come out as like you know. Like if I know I'm gonna open the case then it'll come out as like

they're gonna be put on the registry even if I don't think they should be.

E: Okay, so you don't have any control over what happens.

A: Yeah, but if I deny the case... if we open the case...it gives us our category. But it really doesn't determine if we're opening them or not. (African-American, Female)

In a moment where she is explaining her lack of discretion, Angela takes care to emphasize that there are other spaces in her work where she possesses decision-making power. Workers' descriptions of how they intervened during the investigative period provide insight into how workers who felt their autonomy restricted in one area were able to take charge in another. Keisha recounts:

K: I had a home that I was supposed to substantiate. I should have. The mom had two children; one was developmentally disabled. The other child, she never diagnosed him [he was also impaired] but she homeschooled him the whole time and he's like 16 to 17. This house, literally, I wish I had my phone. I could show you. Trash stacked five feet high in the home; every single room. They had it to where they couldn't even get into their bedrooms and they had made a small path to the bathroom and they were all sleeping on the couch in the living room.

E: So why didn't you substantiate it?

K: 'Cause where was this child gonna go? And it was two children in the home and that boy was 17. So with my supervisor's permission, I stayed in this house for six months. I literally cleaned it up.

E: Yourself?

K: I went and got the dumpster. She stayed in the trailer homes, but she had a doublewide trailer

E: So you kept the investigation open for six...

K: To help her clean it. t took three months to get the trash out of there.

E: And the supervisor didn't know?

K: She knew. I communicated with her. But when she was telling me that I need to do it, I requested for this to stay open a little longer. (African-American, Female)

By all accounts, there is a clear preponderance of evidence to substantiate the mother in this case for physical neglect as the house does not meet the policy for being safe nor does it meet the requirement that children have beds to sleep in.

Keisha possesses ample reason to conclude her investigation and move forward with the case. However, Keisha knows that if she were to fill out the risk assessment at this time, the case would likely be categorized as intensive and require a removal as indicated by her statement that the children do not have anywhere else to go.

Keisha clearly does not want the situation to be out of her control and goes about fixing the problem before the investigative period concludes. Cleaning the place up, Keisha is able to make the situation safe by staging the home so that it does not add points to the SDM score before she finalizes the investigation. In this way, Keisha exercises an extreme amount of discretion and professional autonomy even though she posits that she does not have much control over case trajectories. While this is an extreme example (keeping a case open for that amount of time is in no way normative), Keisha's actions highlight the creative ways that investigators who understand their autonomy to be limited can influence the trajectory of their cases. Gabrielle explains how this process can work in a less extreme way:

> People, you'll find have their own judgments come in to play when they go to people's homes. So that's why I focus on the policy... You could have feces all over the floor if you want to, you know, I mean it shouldn't be like that bad, but I mean they there's no need for removal, in that case. All you do is give the family a few days to clean it up, that's it. (African-American, Female)

During the investigative period, Gabrielle offers families a chance to change the conditions of the complaint. If a family has cleaned up the house by the end of the investigative period, Gabrielle will have a clear rationale to deny cases that could be classified as high risk if they were opened.

The decision to substantiate a person for child abuse or neglect is a powerful one and workers report considering the impact of substantiation on family life when making it. Dorothy told me about her decision-making process regarding whether

to substantiate a charge of medical neglect against two parents. In this case, a cognitively impaired school age child was riding her bike when the back end was clipped by the neighbor's car and she fell. The child's therapist heard the story, spoke to the parents about her concerns, and then reported them to Child Protective Services:

> So she [the therapist] talks to mom. Why didn't you take her to the hospital? You're medically neglecting your child. And so mom is like, well we used our best decision. She didn't have any bruises. The car did not hit her, it hit her bike tire and she fell to the ground. She had a scraped knee. We thought it was the equivalent of your child just falling off of her bike, so we didn't feel the need to take her [to the hospital].

Dorothy went on to explain that during the investigative period she had the child examined, with the doctor finding that no serious injuries resulted from the accident. The only physical effect was bruising on the child's abdomen. Dorothy described her thought process about substantiating this way:

> Like I mean if it was me, personally, if it was my kid, I probably would've taken her to the doctor just to be safe, just because...you know what I mean? You weren't out there and actually saw what happened. The neighbor walked her home and then you found out. But listening to both mom and dad, you know they did what they thought was right. And so I just felt like there's no reason to open this for medical neglect and tarnish their record when they really weren't being malicious about it you know (African-American, Female).

Dorothy has enough cause to substantiate for medical neglect but did not feel that such an action would be justified based on the parent's good intentions and reasoning (that the child had no visible injuries). By choosing not to substantiate, Dorothy actively works to protect the parents from any consequences that could arise from having a substantiated charge of medical neglect. In this way, Dorothy exercises control over the trajectory of her case even though she experiences her discretion as relatively hindered in other areas of the process.

Like Dorothy, Angela uses the investigative period to influence outcomes for cases. Angela detailed a neglect case that she denied when she also had enough evidence to support opening it:

> A: So when I talked to the mom about her history and things like that, she was telling me about her husband being very abusive and she left. He cut the cable wires, he would not let her wear dresses, he didn't want her to work. She was like, you know I left him once before, but he threw a vacuum cleaner at my head and hit me in the head with it and that was the last straw. They had had some history where the 7-year-old had called the police on dad for like fighting with mom. So and I said, well is he paying child support? She said, no, he told me that I get food stamps and that's my child support and that he's not willing to help me out.
>
> So this mom was going through a lot. I mean she's doing the best that she can to provide for her five kids. She has some of the kids in counseling because they weren't coping well with dad not being there. So she's doing everything. You know what I mean? So I felt really confident about denying that case. And what ended up happening is, I said, what do you need? You know anything that I can help you get? She's like, well my kids don't have beds and they had mattresses on the floor. I got them five new beds.
>
> I was trying to be sure that my supervisor understood why, even though yes, the kids don't have coats all the time; okay, yeah maybe the food has poor nutrition. But I'm you know, letting her know what's going on; like how this working mom is trying to be a single mom and be a good mom and trying her best to...some of these kids got poor behavior you know. Like the 12-year-old was like cursing at her. So...[I was] trying to make sure my supervisor understood why I was recommending that it be denied.
>
> E: Do you remember what the risk level was for that case?
>
> A: I think it might have been moderate. (African-American, Female)

While there was enough cause to open a case against this mother, Angela uses the discretion granted her during the investigative period to protect a mother that she does not believe deserves to be in the child protective services system because she has been a victim of abuse and has "good intentions."

These examples indicate that discretion and professional autonomy are not an all or nothing experience for workers with a more restricted understanding of

their decision-making abilities. Debates about the place of practice wisdom, intuitive thinking, clinical judgment, and discretion in social work practice tend to frame these approaches in opposition to evidence-based practices. Critics argue that "social work has been transformed from a self-regulating professional activity into a managed and externally regulated set of tasks" (Evans & Harris, 2004). However, I find that the picture is more complicated in practice. Rather than characterizing discretion as 'all or nothing' or 'good or bad,' discretion should be conceptualized as taking place on a continuum (Evans & Harris, 2004).

I have demonstrated that even workers who do not understand themselves to have autonomy in their casework still find ways to influence the direction of cases in areas that have not prescribed how action can be taken. Conversely, I have also found that workers who experience themselves to have a high degree of autonomy still follow the actuarial based models, even though they may not use them as intended. By conceptualizing discretion on a continuum and by acknowledging that workers make case judgments alongside and within protocoled environments, an understanding of the utilization of standardized decision-making emerges that neither romanticizes the ability of actuarial based risk assessments to eliminate subjective assessments in casework nor overly emphasizes the role that discretion plays in the context of structured decision-making models.

Privilege, Power, and Protocols

Distinct patterns emerge when examining which workers felt that they had autonomy over their decision-making and which workers felt more restricted. Workers who articulated a strong sense of professional autonomy and discretion tended to be white men and/or workers with extensive experience at CPS. In

contrast, workers with a restricted sense of professional autonomy and discretion tended to be women who are racial and ethnic minorities, men and women of any ethnicity who have 0-2 years of experience at CPS, and/or women for whom maintaining order is a central part of their work professional identity.

How are we to understand these patterns in the context of a highly structured decision-making environment? These patterns suggest that the Structured Decision-Making Model assessments are not neutral objects; rather their use is socially situated as well as relationally contingent. An intervention that is intended to reduce the impact of individual positionality cannot and does not fully blunt the force of social power and privilege issues that are at play in everyday life.

Sociological research has consistently demonstrated that power and privilege mediate and inform approaches to the world. People in positions of power and privilege often look at the same situation differently than those in less privileged positions. For example, Lareau (2003) finds that social class determines how parents approach the project of raising their children with people from lower socio-economic backgrounds articulating a strategy of natural growth and parents from middle and upper socio-economic engaging in "concerted cultivation" where they actively structure their family life to influence achievement. Bell (2014) compares class experiences of infertility demonstrating that women of lower socioeconomic status want pregnancy to be a natural process compared to women of higher socio-economic status who engage in technological solutions to achieve pregnancy. In each of these examples, social class shapes both the perception of and response to a particular phenomenon. Numerous sociological investigations confirm that power and privilege in the form or race, gender, class, and sexual orientation inform how

specific situations are conceptualized and authored (Allan, 2006; England, McClinktock, & Shafer, 2011; Klinenberg, 2003; Martin, 1998; Pugh, 2009).

A sociological view of power and privilege informs our understanding of differential conceptions and uses of the SDM on the ground. While one of the primary goals of actuarial based assessments is to even the playing field for both workers and families by neutralizing subjective interpretations of specific situations through the administration of an objective inventory, my findings raise questions about the ability of these assessments to do just that. It is clear that the participants in this study each bring their own views of the SDM to bear on their work and that these views, in turn, shape how they respond to structured decision-making aids. In this way, power and privilege shapes how workers in distinct status groups use actuarial decision-making aids differently.

A Shield or A Sword?

Autonomous workers tended to use the SDM as a sword through which to represent their own views of a case. If the SDM did not reflect their own interpretation of the facts of the case, these workers would either reinterpret questions on the SDM to change the outcome or "finagle" their response to ensure that the official record reflected their personal and professional judgments. Autonomous workers possessed some permutation of privilege in the workplace whether based on race, gender, or seniority or some combination of these social categories. Workers who possessed more institutional or cultural power found maintaining their autonomy and discretion to be a particularly salient aspect of preserving meaning in their professional work. As Espeland (1997) in a review of Porter's Trust in Numbers writes: "Savvy bureaucrats know that even when

decisions are obvious, it is often far better to send superiors a range of plans to choose from since displays of discretion are as important as symbols of power as they are symptoms of power" (Espeland, 1997). Here, the display of discretion and options is not necessarily for the benefit of a bureaucratic supervisor but for the worker himself. Men in the field of social work are often underrepresented and the lack of prestige that is attached to the position may make demonstrations of authority more important in a field where masculinity, intelligence, and expertise are questioned.

The performance of discretion is also the performance of authority. It may be that autonomous workers not only felt more entitled to exercise their judgment but that they were the only group that felt entitled to make a judgment at all. Workers in this group tended to frame subverting the findings of the SDM as the only moral choice given the high stakes nature of their work. Opposition to the SDM, in cases, they believed the SDM to be incorrect was presented as the ethical thing to do. Workers with either high levels of experience or those who were men trusted their judgments to begin with and did not feel that they needed the assistance of a tool, particularly one that they did not trust. These findings confirm tropes of gender in the workplace, namely that men feel more entitled to make decisions and pursue autonomous practices. In this case, those workers that felt more entitled to having an opinion were more likely to rely on their gut or intuitive reasoning than workers who occupied positions of lower social status, particularly were gender was concerned.

Many workers who reported having a limited sense of professional autonomy and discretion used the SDM as a shield for their decisions. Positively restricted

workers described feeling that the SDM provided a framework from which to make objective judgments that protected the families they were evaluating from their own personal biases. Because the SDM was based in scientific fact, these workers saw the SDM as providing cover from the enormous responsibility that comes with making decisions about whether children are safe to remain in their homes or whether they must be removed. The SDM and a more limited sense of professional autonomy also protect workers from experiencing the profound consequences of being "wrong."

Encouraged to both minimize contact with the legal system as well as make accurate assessments of risk that err on the side of caution, child welfare workers face frequent, countervailing pressures as they formulate their recommendations for families. It is no surprise then that more inexperienced workers (workers with less than two years experience) might place an enormous amount of faith in the SDM or be more accepting of the limits that following a standardized protocol places on their decision-making capacities. This is one way to understand why the only white man in my sample embraced the SDM, as he had less than one-year experience on the job. In this case, being a new worker might be a more salient than identification with a more powerful group based on race and gender.

More inexperienced workers, or workers who were in less powerful social positions based on race and gender, also used the SDM as a shield when challenging decisions that supervisors made and with which they disagreed. Restricted workers often used the SDM to advocate for the closure of a case that they believed was being kept open unfairly. The SDM provided a space through which they could challenge a case's trajectory without necessarily directly challenging their supervisor's expertise or judgment. Often this was a successful strategy that not

only facilitated the changes that a worker had desired but preserved the sense that a worker was cooperative rather than "difficult"--a label that proves particularly threatening to a workers professional trajectory in a highly bureaucratic organization.

Racial and ethnic minorities may have been more likely to follow the SDM or see the SDM as a positive tool because they have either experienced bias themselves, which was detrimental in many aspects of their personal lives or because they saw the SDM as protecting them against accusations of bias or unprofessionalism. Aware of the role that bias can play in decision-making, racial and ethnic minorities might be more willing to accept the trade off of more limited professional discretion and authority for a system that they may have viewed as more fair or just and which they hoped would protect themselves and their clients from discrimination, ignorance, or bias.

Or racial and ethnic minorities might understand the SDM and the more limited sense of professional autonomy that accompanies its use as important to follow because their work and decision-making was more likely to be scrutinized and questioned. Racial and ethnic minorities must find a way to continually demonstrate their capacity and professionalism when their abilities might presumed to be less than other workers. Closely adhering to the SDM and other departmental policy is a formal and routinized way to display one's competence and professionalism in a manner that is difficult to dispute. Racial and ethnic minorities that disagree with the SDM or do not see it as a valid and reliable professional tool might be more willing to accept the limitations on their professional autonomy and discretion because they either feel more closely examined in the workplace or

because they believe that they have more to loose than other workers who feel more comfortable with answering the SDM in a way that reflects decisions that they have already made. Whichever the case, following protocols and policy was a deeply salient part of the experience of being a CPS worker for this group.

Incorporating Reflective Practices Into Standardized Assessments

As this paper demonstrates, even within a highly structured and standardized environment there exists a wide range of responses to and utilization of an actuarial based risk-assessment. How agentic a worker feels in her professional role combined with a worker's stance on the validity of the SDM, impacts how she will integrate the model into her actual decision-making. Workers, who understand themselves to have professional authority and discretion, report making the risk assessment fit their case decision-making rather than basing their case decision-making off of the risk assessment. These workers are able to satisfy the formal requirements of state policy without actually utilizing the intervention.

Restricted workers take special care to complete the SDM as intended but even within this group there is no uniform opinion about whether the SDM is effective, valid, reliable, and fair. While restricted workers attempt to engage with the intervention in the "right" way, the intervention itself may not be able to completely control for workers subjective ideas about the case either. First, the same subjective processes that workers with discretion employ to alter the outcome of the SDM might be occurring here as well but without conscious intent. Second, workers in this category deploy other strategies for maintaining autonomy in their casework, such as using the investigative period to affect outcomes that become obscured when

one only looks at how these workers recount they use the SDM in their everyday work.

Thinking about how the intersection of years of experience, race, and gender influences workers' response to the SDM invites a deeper exploration of the meaning of cultural processes in Implementation Science. When researchers and practitioners call for changes in workplace culture to ensure that interventions are effective, we need to think more explicitly and more carefully about the meaning of cultural change. Whose cultures are we changing? Whose viewpoints are we privileging? Which ideas and from which workers are we bringing to the forefront? Which normative frameworks are being called into action and which are being left behind? Unpacking the construction of interventions is a key and often overlooked task in policy, model development, and evaluation research. Yet, the normative frameworks on which interventions are constructed and how they are then responded to are likely critical factors in how an intervention is performed, integrated, or received by the population at which it is aimed.

Here, power plays a very real role in the construction, adoption, and deployment of an evidence-based intervention. In the data presented above, workers who occupy what have been traditionally more marginalized positions in society are the ones attempting to ensure that their work has "high fidelity" to the SDM model, while workers in more privileged positions consciously choose to subvert it. However, it remains unclear whether workers who occupy more marginalized positions would act in similar ways to their colleagues if they felt that they had the power to do so. Therefore, it is not evident that simply changing the culture of more experienced workers and making them act more like less experienced ones

would really be sufficient to improve implementation of the SDM, especially given that one of the issues at play is workers belief in the utility of the intervention itself. In this way, cultural, individual, and intervention related factors collide in messy ways to challenge traditional thinking about the implementation of interventions. These findings suggest that workers' positionality, social standing, cultural location, and individual subjectivity cannot entirely be eliminated from practice innovations meant to contain them.

At the same time, these findings emphasize how important cultural factors are in the deployment and experience of interventions. Social demographics, years of experience, power, and attitudes about work clearly shape individual responses to the SDM, patterning both attitudes about effectiveness and how workers apply the SDM to their decision-making. Normative frameworks about what work means, what being a professional means, what it means to be a "good employee," what it means to be fair, and what is safe for a child, all play out in relationship to how workers choose to approach the SDM. While the SDM was developed, in part, to respond to how diverse ideas are about each of these concepts, it is not wholly successful in being able to always control for, objectify, or flatten workers individual views on these ideas. Subjective assessments, which were, in part, rooted in cultural locations were always present in workers application of the evidence to the risk assessment.

Common suggestions to improve the implementation of interventions such as changing the culture of the office to increase compliance with the SDM; creating a highly structured environment within which the SDM is deployed (as State A has already done); introducing implementation strategies to improve individual

performance such as more detailed education about how to perform a model; or increasing oversight to improve model fidelity are unlikely to be effective strategies when there are such a diversity of approaches to the SDM. Individual ideas, locations, and subjectivities produce strong viewpoints that impact an intervention's implementation and which must be addressed when there is a large gap between how an intervention is intended to be performed and how it is being performed in practice.

Introducing opportunities for workers to reflect on how their own viewpoints intersect or challenge the findings of the SDM might be a more successful strategy. Dealing directly with the unconscious or conscious beliefs that every individual has about cases in the general and in the specific, creates a space for actively engaging what workers now either find a way to subvert or ignore. Instead of assuming that we can develop new and better methods for eliminating subjective assessments and processes, we should assume that subjective assessments and processes are not only an essential component of human experience and interaction but an important part of what many more experienced workers understand to be a central part of what makes them seasoned professionals. Rather than engaging in a tug of war between control and subversion of these subjectivities, it makes more sense to engage them openly on the assumption that they are in fact ineliminable.

Throughout my fieldwork I find many examples of instances where workers employ their subjective assessment to combat what seem like illogical decisions made in a rote or punitive systematic context. Simultaneously, I observed other examples were workers decisions to act outside policy based on their subjective assessment of a case were simply quite terrifying to me. A catch all approach that

conceptualizes subjectivity and discretion in decision-making in binary terms (bad or good) misses three key points (1) that discretion can be a positive or negative practice depending on the case, the worker, and the organizational environment; (2) policies that attempt to either limit or encourage discretion must be effective for the full spectrum of workers from the most inexperienced and unsophisticated to the most practiced and wise; and (3) even in a highly structured and protocoled work environment, workers find a way to maintain some level of authority and discretion in their casework. Power struggles about how much of workers ideas or knowledge can play a role in decision-making are likely to have a Wile E. Coyote and Road Runner feel to them--an endless chase where the terms may change but the outcome stays relatively similar, if not the same.

These power struggles echo epistemological debates about whether social work is an art or a science continuously play out with little resolution. It is clear that evidence-based practice and more standardized policies and procedures are here to stay. However, more standardized procedures do not have to mean, as it does now, that reflective practices, clinical judgments and assessments need to be left by the wayside.

Reflective practices could be incorporated into State A's current Child Protective Service policy by requiring a narrative response to the risk assessment by workers; review and discussion of each case with an experienced clinician; and opportunities to raise or lower the risk assessment after thorough case conferences. These actions would create a system to address, upfront; cases were there is a tension between the formal policy and workers findings. Airing these differences in the open and then actively engaging them in the context of a layered or team review

creates a system for taking seriously and being respectful of the professional autonomy and training of CPS workers. Actively seeking worker's thoughts about a case is likely to spur a more honest and open conversation among all workers and could do much to even out some of the power differentials based on worker positionality described in this chapter.

At the same time, incorporating reflective practices as a part of standardized work will be useful to the entire spectrum of workers that CPS employs. Clinical case reviews and reflective narratives both enable the most sophisticated workers to challenge policy that does not make sense for a specific case while also having the potential to catch subversions of policy that are based on strong transference or countertransference of the worker and which are likely inappropriate given the facts of the case.

Works Cited

Allan, J. (2006). *Privilege, Power, and Difference*. New York: McGraw-Hill.

Baird, C., & Wagner, D. (2000). The relative validity of actuarial and consensus based risk assessment systems. *Children and Youth Services Review, 22*, 839-871.

Baird, C., Wagner, D., Healy, T., & Johnson, K. (1999). Risk Assessment in Child Protective Services: Consensus and Actuarial Model Reliability. *Child Welfare, 78*(6), 723-748.

Barnett, D., Manly, J., & Cicchetti, D. (1993). Defining Child Maltreatment: The Interface Between Policy and Research. In D. Cicchetti & S. Toth (Eds.), *Child abuse, child development, and social policy*. Norwood, NJ: Ablex.

Charmaz, K. (2006). *Constructing grounded theory: a practical guide through qualitative analysis*. Thousand Oaks: Sage.

Coulton, C., Corbin, J., & Su, M. (1995). Community level factors and child maltreatment rates. *Child Development, 66*, 1262.

Davidson-Arad, B., & Benbenishty, R. (2010). Contribution of child protection workers' attitudes to their risk assessments and intervention recommendations: a study in Israel. *Health & Social Care in the Community, 18*, 1-9.

Dawes, R. M. (1988). *Rational choice in an uncertain world*. Orlando: Harcourt, Brace, Jovanovich.

Dawes, R. M. (1999). A message from psychologists to economists: mere predictability doesn't matter like it should (without a good story appended to it). . *Journal of Economic Behavior & Organization, 39*(1), 29-40.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical vs. actuarial judgment. *Science, 243*, 1668-1674.

DePanfilis, D. (1996). Implementing child mistreatment risk assessment systems: lessons from theory. *Administration in Social Work, 20*(2), 41-59.

Dettlaff, A. J., Rivaux, S. L., Baumann, D. J., Fluke, J. D., Rycraft, J. R., & James, J. (2011). Disentangling substantiation: The influence of race, income, and risk on the substantiation decision in child welfare. *Children and Youth Services Review, 33*(9), 1630-1637. doi: 10.1016/j.childyouth.2011.04.005

Dettlaff, A. J., & Rycraft, J. R. (2008). Deconstructing Disproportionality: Views From Multiple Community Stakeholders. *Child Welfare, 87*(2), 37-58.

Dettlaff, A. J., & Rycraft, J. R. (2010). Factors Contributing to Disproportionality in the Child Welfare System: Views from the Legal Community. *Social work, 55*(3), 213-224.

Dorsey, S., Mustillo, S., Farmer, E., & Elbogen, E. (2008). Caseworker assessments of risk for recurrent maltreatment: Association with case-specific factors and re-reports. *Child Abuse & Neglect, 32*, 377-391.

Dworsky, A., White, C. R., O'Brien, K., Pecora, P., Courtney, M., Kessler, R., . . . Hwang, I. (2010). Racial and ethnic differences in the outcomes of former foster youth. *Children and Youth Services Review, 32*(6), 902-912. doi: 10.1016/j.childyouth.2010.03.001

Emerson, R., Fretz, R., & Shaw, L. (1995). *Writing Ethnographic Fieldnotes.* Chicago: University of Chicago Press.
England, P., McClinktock, E., & Shafer, E. (2011). Birth control use and early, unintended births: Evidence for a class gradient. In M. Carlson & P. England (Eds.), *Social Class and Changing Families in an Unequal America.* Stanford: Stanford University Press.
English, D., & Pecora, P. (1994). Risk assessment as a practice method in child protective services. *Child Welfare, 73,* 451-473.
Espeland, W. (1997). Authority by the numbers: Porter on quantification, discretion, and the legitimation of expertise. *Law and Social Inquiry, 22.*
Evans, T., & Harris, J. (2004). Street Level Bureaucracy, Social Work and the (Exaggerated) Death of Discretion. *British Journal of Social Work, 34,* 871-895.
Fluke, J., Yuan, Y., Hedderson, J., & Curtis, P. (2003). Disproportionate representation of race and ethnicity in child maltreatment: Investigation and victimization. *Children and Youth Services Review, 25.*
Gambrill, E. (1990). *Critical thinking in clinical practice: improving the accuracy of judgments and decisions about clients.* San Francisco: Jossey-Bass.
Gambrill, E., & Shlonsky, A. (2000). risk assessment in context. *Children and Youth Services Review, 22*(11-12), 813-837.
Gambrill, E., & Shlonsky, A. (2001). The need for comprehensive risk management systems in child welfare. *Children and Youth Services Review, 23*(1), 79-107.
Giovannoni, J., & Becerra, R. (1970). Defining Child Abuse. New York: Free Press.
Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical statistical controversy. *Psychology, Public Policy, and Law, 2*(293-323).
Holland, S. (2000). The assessment relationship: interactions between social workers and parents in child protective assessments. *British Journal of Social Work, 30,* 149-163.
Holland, S. (2001). Representing Children in Child Protection Assessments. *Childhood, 8*(3), 322-339.
Hutchinson, E. (1990). Child Maltreatment: Can it be defined? *Child Welfare, 49*(4).
Jones, L. (1993). Decision making in child welfare: A critical review of the literature. *Child and Adolescent Social Work Journal, 10*(3).
Klinenberg, E. (2003). *Heatwave: A social autopsy of disaster in chicago.* Chicago: University of Chicago Press.
Lindsey, D. (1992). Reliability of the foster care placement decision: A review. *Research on Social Work Practice, 2*(1), 65-80.
Loman, A., & Siegel, G. (2004). An Evaluation of the Minnesota SDM Family risk assessment. St. Louis, Missouri: Institute of Applied Research.
Martin, K. (1998). Becoming a Gendered Body: The practices of preschool. *American Sociological Review, 63*(4).
McDonald, T., & Marks, J. (1991). A Review of risk factors assessed in child protective services. *Social Service Review, 65*(1), 112-132.
Munro, E. (1996). Avoidable and unavoidable mistakes in child protection work. *British Journal of Social Work, 26,* 793-808.

Munro, E. (2004). A simpler way to understand the results of risk assessments. *Children and Youth Services Review, 26*, 873-883.
Osmo, R., & Benbenishty, R. (2004). Children at risk: rationales for risk assessments and interventions. *Children and Youth Services Review, 26*, 1155-1173.
Portwood, S. (1998). The impact of individuals' characteristics and experiences on their definitions of child maltreatmen. *Child Abuse & Neglect, 22*(5).
Pugh, A. (2009). *Longing and Belonging: Parents, children and consumer culture.* Berkely: University of California Press.
Rose, S. J. (1999). Reaching Consensus on Child Neglect: African American Mothers and Child Welfare Workers. *Children and Youth Services Review, 21*(6).
Rose, S. J., & Meezan, W. (1996). Variations in perceptions of child neglect. *Child Welfare: Journal of Policy, Practice, and Program, 75*(2), 139-160.
Schwalbe, C. (2004). Re-visioning risk assessment for human service decision making. *Children and Youth Services Review, 26*, 561-576.
The Structured Decision Making Model: An Evidence-based Approach to Human Services. (2008). Madison: Children's Research Center.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainity: hueristics and biases. *Science, 185*, 1124-1131.
Tversky, A., & Kahneman, D. (2002). Extensional vs. intuitive reasoning: the conjunction fallacy in probabilistic judgment. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Hueristics and biases: the psychology of intuitive judgment.* New York: Cambridge University Press.
Wald, M. S., & Woolverton, M. (1990). Risk assessment: The emperor's new clothes? *Child Welfare, 69*(9), 483-511.
Walsh, K., Bridgstock, R., Farrell, A., Rassafiani, M., & Schweitzer, R. (2008). Case, teacher and school characteristics influencing teachers detection and reporting of child physical abuse and neglect: Results from an Australian survey. *Child Abuse & Neglect, 32*(10), 983-993. doi: 10.1016/j.chiabu.2008.03.002
Weiss, R. (1995). *Learning From Strangers: The Art and Method of Qualitative Interview Studies.* New York: The Free Press.
Young, L. (1964). *Wednesday's children: A study of child neglect and abuse.* New York: McGraw-Hill.
Zuvarin, S. J. (1999). Child neglect: A review of definistions and meausrement in research. In H. Dubowtiz (Ed.), *Neglected children: Research, practice, and policy.* Thousand Oaks: Sage.

Chapter 3
The Goldilocks Problem: The Local Context of Standardization

"On the ground, every standard is simultaneously over determined and incomplete."
Timmermans and Epstein, 2010

Introduction

In the last decade, Evidence-Based Practice (EBP) has become the new paradigm for social work research, practice, and education in the United States (Adams, Matto, & LeCroy, 2009; E. Gambrill, 1999, 2001; Eileen Gambrill, 2010). Derived from the Evidence-Based Medicine movement (EBM) and paralleling similar movements in psychology, education, nursing, and criminal justice, EBP seeks to offer a systematic and empirical way for making decisions about which treatments and interventions to use with social work clients (Eddy, 2005). The goals of the Evidence-Based Practice Movement are to ground social work practice in empirically tested interventions, to standardize a framework for selecting interventions, to "demonstrate [e] that treatment decisions are transparent and justifiable," and to provide evidence of the cost effectiveness of interventions (Gambrill, 2006,). Embedded within these goals is the idea that an intervention can and should be standardized across contexts. While debates about the epistemology, goals, use, and merits of EBP in Social Work abound, these debates tend to focus on whether systematized ways of knowing should trump practice-based wisdom or

clinical and professional judgment (Atherton, 1993; Brekke, 1986; Fischer, 1973; Franklin, 1995; Fraser, Taylor, Jackson, & Ojack, 1991; E. Gambrill, 1999, 2001; E. D. Gambrill, 2002; Gibbs & Gambrill, 2002; Ivanoff, Blythe, & Briar, 1987; Longhofer & Floersch, 2012; Peile, 1988; Staller, 2006; Thyer, 1994; Wakefield & Kirk, 1996; Witkin, 1996; Witkin & Harrison, 2001) or on the limitations and complications that accompany the implementation and translation of Evidence-Based Practice in real world settings (Adams, et al., 2009; Ammerman, Smith, & Calancie, 2014; Berger, 2010; Bergstrom, 2008; Brekke, Ell, & Palinkas, 2007; Chorpita, Becker, & Daleiden, 2007; Flanagan, 2013; Green, 2006; Longhofer & Floersch, 2012; Mullen, 2014; Palinkas & Soydan, 2012; Parsonson, 2012).

Whether framed within ideological or practical terms, discussions about the utility of standardized work products often accept the premise that these products actually *do* create uniformity in work practices. The assumption of standardization granted, disagreement centers on whether uniformity facilitates or impedes good work, whether standardization is a value neutral or value laden process that can lead to objective results, and how standardization could be improved to create 'better fidelity to the model' on the ground. While these are important debates, this line of intellectual inquiry elides a more basic question: *Can we expect standardized tools to function as intended across practice settings?*

This paper examines the deployment of a standardized procedure in two different states to investigate whether a standardized tool for child welfare decision-making equalizes performance across contexts. Demonstrating that the local policy and administrative context of each state structures the meaning and use of the tool by child welfare workers, I argue that standardization is both a social and a technical

process. This work contributes to a growing literature in implementation science that calls for contextual concerns to be a primary part of translational science and challenges the idea that simply engaging in better training of a particular model will necessarily improve the deployment of an intervention. The predominance of standardized interventions as a paradigm for social work intervention at both a policy and practice level makes an investigation of whether standardized practices actually create practice uniformity across settings both necessary and timely.

In this paper, I briefly contextualize the Structured Decision-Making model within the Evidence-Based Practice movement. I then problematize the assumption that standardization is a technical process and that lack of fidelity to a model can be addressed only through training and education. Following this, I describe the research study and findings. This paper concludes with a discussion of the implications of this research for the diffusion of interventions and for pursuing adaptive vs. restrictive implementation strategies.

Evidence-Based Practice and The Structured Decision-Making Model

The belief that social work interventions must be based on empirically validated outcomes is widely accepted. Delivering efficacious treatment is a social work value embodied within the professional code of ethics calling for social workers to "base practice on recognized knowledge, including empirically based knowledge." (NASW 1999, sec. 4.01.c). The accreditation body for social work education now requires programs to train their students in evidence-based practices. Clearinghouses such as the such as the National Registry of Evidence-Based Practices and Programs (NREPP, 2007; the Substance Abuse and Mental Health Services Administration) have emerged to assist practitioners in utilizing validated

interventions, while national funding streams for research emphasize the development and translation of empirically based interventions in all aspects of the funding process (Adams, et al., 2009). Not only is EBP considered a critical part of good social work practice and the gold standard for research but the converse is also true: social work interventions that are not evidence-based are understood to be "bad practice." With this understanding, policy-making bodies have begun to link social service provision to Evidence-Based Practices with many states mandating that all state funded interventions be grounded in this model.

The Structured Decision-Making Model (SDM) is one evidence-based practice intervention that has proliferated in the last decade as social work has taken a turn toward empirically tested practices. Currently in use in over 30 states, the SDM is designed to address issues in decision-making among child welfare workers. Research has consistently demonstrated that child welfare decisions are extremely variable, even when workers are considering the same set of facts (Davidson-Arad & Benbenishty, 2010; Lindsey, 1992; Rossi, Schuerman, & Budde, 1999). Developed by the Children's Research Center, the Structured Decision-Making Model is intended to eliminate some of the most glaring difficulties in inconsistent decision-making such as subjectivity, bias, and the lack of stable criteria for case evaluations. Built around a series of assessments including an intake assessment, a safety assessment, a risk assessment and reassessment, a family strengths and needs assessment, a child strengths and needs assessment, and a family reunification assessment, the SDM delineates the key decision-points in the life of a child welfare case into a discrete set of manageable tasks. Grounded in actuarial science, the SDM guides decision-making through probabilistic appraisals based on the presence (or absence) of a

series of well-validated risk factors. In doing so, the SDM transforms decision-making in child welfare from a relatively personal process to a systems-level task. Further, the SDM seeks to reduce the racial disproportionality that comprises and compromises the child welfare system by creating consistent evaluations of families using empirically validated criteria. If each measure applies equally to every family no matter what their race, class, or ethnic background, then the assumption is that families will be evaluated fairly and racial and ethnic disproportionality in the child welfare system that result from unfair evaluations will be decreased. The core principles of the SDM are: (1) reliability; (2) validity; (3) equity; and (4) utility. The table below details these principles as the developers describe:

> **Reliability:** SDM assessments systematically focus on the critical decision points in the life of a case, increasing worker consistency in assessment and case planning. Clients are assessed more objectively, and decision-making is guided by facts rather than individual judgment.
>
> **Validity:** The actuarial research-based risk assessment, which accurately classifies families and clients according to the likelihood of subsequent outcomes, enables agencies to target services to clients at highest risks for negative outcomes, such as maltreatment recurrence or difficulty finding and maintaining employment.
>
> **Equity:** SDM assessments ensure that critical case characteristics, risk factors, and domains of functioning are assessed for every client, every time, regardless of social differences. Detailed definitions for assessment items increase the likelihood that workers assess all clients using a common framework.
>
> **Utility:** The SDM model and its assessments are easy to use and understand. Assessments are designed to focus on critical characteristics that are necessary and relevant to a specific decision point in the life of a case. By focusing on critical case characteristics, workers are able to organize information gathering and case narratives in a meaningful way. Additionally, the assessments facilitate communication between worker and supervisor, and unit to unit of the status of each case.
> (http://www.nccdglobal.org/assessment/structured-decision-making-sdm-model)

Embedded within every principle of the SDM is the concept of standardization. By unifying each aspect of the decision-making process around a set of common criteria or actions, the SDM enables decision-making to be done at two levels: (1) across workers with different sets of experiences and training and (2) across policy environments. The Children's Research Center unequivocally states the overall benefits to a standardized approach, claiming that the SDM is not only the most accurate form of assessment available in child welfare but also a system that will improve child welfare outcomes across a number of dimensions.

In this article, I argue that the Structured Decision-Making model is not only sensitive to organizational contexts but that the organizational context in which it is deployed also structures the SDM's meaning and use. Rather than standardizing work practices across all child welfare workers and policy environments, how workers understand and then utilize the SDM reflects a variety of organizational factors and orientations. As a practice technology, the significance of the SDM becomes socially created in every setting in which it is implemented. In turn, how professionals manage their own subjectivity in the face of new tools like the SDM becomes highly dependent on the meaning these tools possess in a specific environment. By unpacking how informal variation in standardized work practices occurs at the organizational level, this paper frames standardization as adaptive rather than technical work (Dixon-Woods, Bosk, Aveling, Goeschel, & Pronovost, 2011). In doing so, I argue that the process of standardization itself is in no way standard.

Methods

Sample

The data for the present study originates from over 150 hours of semi-structured, qualitative interviews with 66 Child Protective Service Workers (CPS), in two states. Interviewees were identified in multiple ways. First, I recruited an initial sample through 3 continuing education trainings for child welfare workers held at a school of social work. At these trainings, I introduced myself, described the study and passed around a sign-up sheet for study volunteers. I then scheduled and conducted interviews with those who indicated interest. Additionally, area directors introduced my research to their staffs and passed around a sign-up sheet at several staff meetings. Following a similar procedure, I contacted volunteers and scheduled interviews with those who demonstrated interest. At these area offices, I also recruited participants using a snowball sample as other CPS workers heard about the study from colleagues and asked to join. Participants in State A were offered a $25 Visa gift card to compensate them for their time while state regulations in State B prohibited me from offering an interview incentive. States have been de-identified to protect participant confidentiality.

Participants

Interviewees ranged in age from 24-62. Participants were comprised of 50 women, and 16 men. Workers interviewed had a variety of experience. 16 had worked in their current position less than 2 years, 11worked in their current position between 2-5 years, 7 worked in their current position between 5-8 years,

and 26 had over 8 years of experience. 6 participants declined to note their years of experience on a survey. Workers had a variety of educational experiences. 34 had received a Bachelors degree, 21 held a Masters in Social Work, 7 held a postgraduate degree in a field other than social work, and 4 declined to note their degree. Fields of specialization ranged from criminal justice, to psychology, to sociology, to social work, with the latter representing the majority of the sample. Participants were front line Child Protective Service workers tasked with either making decisions either about substantiation and severity of child maltreatment (Investigators) or, in the case of substantiated cases whether a family was complying with treatment and recommended services (Treatment Workers) or Supervisors who oversaw the decision-making of front line workers. The sample was almost evenly split between workers based in State A (35 out of 66) and State B (31 out of 66).

Interview Content

The interview protocol for this study was designed to assess participants conceptions about their work and how they translated these conceptions into every day work routines (Charmaz, 2006). Semi-structured and open-ended, the protocol begins with a discussion about how workers go through their day. I began the interview by asking how the participant came to child welfare work. Following this initial discussion, I asked participants "to walk me through" the process in a neglect case from the time they received it to when they transferred or closed it or in the case of supervisors "to walk me through how they supervise a case from the time a worker receives it until a disposition has been made"(Weiss, 1995). I collected

information on recent cases, challenging cases, and cases where workers felt confident in their decision-making. These questions allowed me to collect information on what workers 'typically' did and why as well as how their actions unfolded in relationship to each case they described. These descriptions gave me a framework for understanding what the everyday work routines were for individual workers in aggregate and how these routines varied based on the specifics of individual cases.

After collecting these accounts, I then asked workers directly about the Structured Decision-Making (SDM) tools (if they had not yet mentioned them at other points in the interview). These questions were also procedural in nature. I wanted to know the how not the why or what during this part of the conversation. To ground our discussion in fact, I generally asked workers to walk me through the Safety, Risk, or Reunification Assessments for the cases they described. As they did so, I was able to collect data on how workers used the Structured Decision-Making model in practice. After getting the how, I asked about the why, which enabled me to collect data on workers attitudes and beliefs about the SDM as well as the weight they gave them in their decision-making.

Throughout the interview, I followed up every question with probes specific to the conversation. Common probes focused around further elucidating the logic at decision-making points and case detail. Interviews ranged from 60-180 minutes with a median interview time of 90-100 minutes. Interviews were audio taped and transcribed. I conducted interviews at neutral locations such as local libraries and a community center with offices available for sign-out. Interviews also took place in

an available room at area Child Protective Service offices. All names of offices, agencies, and people are pseudonyms .

Data Analysis

In keeping with grounded theory, I began data analysis at the same time that I was collecting data. (Charmaz, 2006). Using an open coding strategy and aided by NVivo 10 qualitative data analysis software, I coded interviews without a particular schema to see what themes emerged (Emerson, Fretz, & Shaw, 1995). I developed concepts and categories reflecting recurring themes in the data and then made sense of how these concepts and categories connected across interviews. After initial open coding, two Research Assistants analyzed the data deductively based on the themes and categories I initially identified. The data presented below represents larger themes found in the data, except where noted. Quotes from the data are edited for clarity (e.g. repeated "Ums" removed), but no words have been changed or re-ordered.

Why Compare States?

State Aand State B offer ideal sites for investigating whether a standardized process achieves a standardizing function. Currently, State A's legislature mandates the use of the Structured Decision-Making Model for child welfare decision-making in all casework. In addition to the SDM, State A has adopted a category system where Child Protective Service workers are required to rank a case based on level of severity. The ranking system further standardizes caseworker action and discretion by specifying the specific protocols that follow each ranking.[1] Case categories are explicitly linked to the risk assessment (RA) with scores on the RA determining the

category a case is assigned. In contrast, while State B does use the SDM model to guide decision-making, the state legislature has not mandated its use. Rather, state child welfare policy incorporates the SDM into the decision-making process without explicitly requiring decisions to be tied to it. Comparing the utilization of the SDM in State A's tightly controlled decision-making environment to State B's fairly open system for determining case actions allows for a deeper investigation into how standardized processes unfold on the ground. Taken together, these differences create rich sources for analysis of how and when standardized processes interact with organizational and policy differences.

The SDM in State A: A Formula for Decision-Making

State A Child Welfare Policy

Structured Decision-Making in State A intends to standardize all points in the child welfare process, establishing the context for a tightly controlled decision-making environment. Mandated for use in child welfare decision-making by the state legislature, the SDM is **the** method by which Child Protective Service (CPS) arrive at their decision. The legislature's codification of the SDM from a suggested practice into law can be understood as an example of the tight integration of EBP with policy in order to transform practice. Policymakers demand that actions taken by bureaucrats be sanctioned by scientific evidence and to which state agencies can be held accountable. The SDM answers that call.

State A has further formalized child welfare decision-making by adopting a category system for classifying child abuse and neglect cases. Rankings standardize caseworker action by assigning specific protocols to each category, which are outlined in the chart below:

State A Child Protective Services Category System

Category V	Following a field investigation, the department determines that there is no evidence of child maltreatment or is unable to locate the family. **No action taken**.
Category IV	Cases where there is no preponderance of evidence that neglect or abuse has occurred. If the risk assessment demonstrates any risk **Voluntary Community Services are Recommended.**
Category III	A preponderance of evidence that child maltreatment has occurred. The risk assessment scores as low or moderate. If the family does not voluntarily participate in services or does not progress CPS may reclassify as a Category II. **Community Services are Needed.**
Category II	A preponderance of evidence that abuse or neglect has occurred. There is a high level of risk for future harm as indicated by the risk assessment. The perpetrator is listed on the Central Registry for Abuse and Neglect. **Child Protective Services are Required in Conjunction with Community Services.**
Category I	A preponderance of evidence that child maltreatment has occurred with the risk assessment scoring as high or intensive. Child Protective Services determine that a court petition is either needed or required. Perpetrator is listed on the Central Registry for Abuse and Neglect. **CPS petitions for the immediate removal of child(ren).Services Provided by Child Protective Services or Foster Care in Conjunction with Community Services.**

Working within this framework, professional autonomy becomes more limited as child welfare workers follow precise directives for case actions according to the assigned category. To determine the category for a case two things are required: (1) a preponderance of evidence (a 51% chance or more that child abuse or neglect has transpired) and (2) the risk level which is established by the score on risk assessment, a component of the SDM. This policy in and of itself requires the SDM to be a critical part of the investigative process and grants the risk assessment formal decision-making power for how a case will unfold.

A Formula For Decision-Making

Workers in State A consistently reported that the SDM does indeed drive formal and informal decision-making practices. The majority of workers (approximately 2/3) described utilizing the risk assessment to establish case actions such as whether a child needs to be removed from the home; the level of service provision a family requires; whether a caregiver should be placed on the Central Registry of Abuse and Neglect; and the level of severity for a particular case. The distinct rules and procedures that link the SDM to case trajectories create an organizational environment that makes it difficult for workers to use any other method for making case judgments. Early in each interview, I asked child welfare workers to describe the general process for neglect cases. A typical account discussed the required response time, the people questioned during the course of their investigation, and their engagement with the Structured Decision-Making model to determine case trajectories. Casey, an investigator characterizes a routine neglect investigation this way:

> E: Sure. So I was wondering if you could tell me, walk me through the process in neglect cases from the time the family is referred until you dispo the case?
>
> C: Okay. Well they'll get referred, usually neglect is considered a Priority 3 case so we have 24 hours to commence it, which just means to find somebody that might have knowledge to whether or not the children are in immediate risk or not. And then you have 72 hours to make face to face with the children so usually I don't see the kids until the next day. I usually don't wait 3 days unless I'm super busy. And then after I interview all the... I'll interview all the children and mom and dad or whatever household members there are. And if dad doesn't live in the house then I have to find dad and go interview him.
>
> And once I've interviewed everybody I conduct a risk and safety assessment to try and figure out what the needs of the family are and what the risk levels are and then we try to address those needs. And as long as they're not imminent risk we'll just refer them to services, either voluntary if it's low risk or tell them we're

> going to get court involved if you don't do these things. And then if it's high risk then we obviously would start filing a petition with court proceedings.

Casey succinctly explains the typical way that a neglect case unfolds. The SDM as symbolized by the risk and safety assessment form an ordinary part of Casey's work process and she relies on the SDM to make case judgments, matching the stated policy goals to the actions she takes. Debra, another investigator, offers a similar perspective:

> Okay. Yep, like we receive a complaint from out centralized intake unit so if we receive it like in the beginning of the day we, depending on what type of complaint it is like a P1, P2 or a P3 depends on how quickly we get to it. But typically you are supposed to commence the case or make contact with somebody that has knowledge of the wellbeing of the child within either 12 to 24 hours. And then you need to see the child within 24 to 72 hours face to face.
>
> So typically, you call the reporting person back, you get as much information as you can. And then if you can't call like a teacher or someone else in the community that knows the child's wellbeing, you need to go out on the case immediately. Make face to face with the victim; interview 'em. And then you interview the biological parents, any like living together partners, relatives that may know information, community members that may have contact with teachers, you know counselors things like that may have contact with the family.
>
> And I guess once you interview all those people and kind of get all the evidence, medical records, police records things like that, then you make a determination whether abuse or neglect occurred. If there is a preponderance of evidence you determine what category it falls into. Like Category 3 is you either open or close with just like a service provider referral or you can open it and monitor the case. Cat 2 is a higher intensity risk level where it needs to be monitored by either the original CPS worker or ongoing worker and then Cat 1 is where they'd be removed from the home and placed into foster care. Then you dispo your case, you close it out depending on what category it is. It may take a couple of months to close out the 2's and 3's. If it's denial it's closed out within 30 days and then on to the next case.

In this long account, Debra uses authoritative language to describe the way she makes her case decisions e.g. *"You determine the category"* and it is unclear whether the SDM plays any role in shaping the trajectory of a case. In order to assess how much weight (if any) Debra gave the SDM in this process, I asked her to walk me

through a case where it was difficult to resolve which category to assign. Debra answered not with a detailed description of a specific case but of the risk assessment itself:

> But with the risk assessment, it's pretty much self-explanatory. When you do the risk assessment that determines the category. Well, it determines between a 2 and a 3, I mean 1 is legal involvement and you choose that and 4 and 5 are denials. So between 2 and a 3 is the risk assessment determines that, so as long as you objectively answer the questions that the risk assessment is asking you, when you answer 'em, then I usually don't question what the risk assessment says.

The active decision-making process that Debra describes is in fact a rather technical task: that of filling out a risk assessment and then using that score to render case judgments. The formal and informal narratives for decision-making have merged to form Debra's process. Keisha, also an investigator, exemplifies how formal policy in this tightly controlled environment can come to structure not only how workers proceed with their case but how they conceptualize it:

> E: Okay. When you're going about making decisions about what services to use, what level to substantiate at, sort of what helps you figure out what to do?
>
> K: Our risk assessment does. It goes on like if there's disabilities in the home; and that's including the kids and the mom; previous CPS involvement including Mom 'cause it could've been that Mom was a product of CPS when she was young, which is a high risk. Because if she was used to that then she doesn't know anything different and 9 times out of 10 she's gonna keep that behavior going with her own children. So that's a major fact. It all depends on how many children are in the home and what's the ages. And within those ages are any children with any disabilities or anything. So all of that plays a key role.

Ann offers a similar assessment

> And the safety and the risk I think play a big role. Because sure they're telling you everything, but it's not until you really go through and you're like reading because they go like, you know they're numbered questions, so when you're going through you're like; yep that applies and yeah that applies and that applies. And it's like okay, this stuff is kind of adding up now. And you know when you're going through the risk because they're asking you about domestic violence and substance abuse and you know past

> history and discipline; and when you're going through there you're like; yeah the child said that, and oh yeah, the child said that; and you know mom admitted to previous DV. So as you're going through there again, it helps you really put things into perspective of; wow this is a lot more than what I had thought or; oh, it's not quite as bad as what I had thought it was going to be. So I mean those really help you factor in.

Both Ann and Keisha identify the items on the risk assessment as the major factors in their decision-making. Formal policy becomes informal practice as Keisha views the items on the risk assessment as not merely probabilistic indicators but as determinative behaviors.

The Safety Assessment is another place where case thinking becomes difficult to separate from formalized case actions. The designation on the Safety Assessment (based on a numerical score of the safety factors present) often acts as a proxy for whether there is a preponderance of evidence in a case. As Justine notes, in doing so, the Safety Assessment then determines whether a case is opened or closed:

> The safety assessment is, I would say, I don't know, I think it's like 15 questions I believe and if there's any safety factors present. If you click, let's say it's a denial and you are saying that there's a safety risk or a safety factor present, I mean you have to basically, there's usually a preponderance if there's a safety factor. So if there isn't a preponderance, there is no safety factor present. If there is a safety factor present then you have to put like, what you are going to do…

This description of the Safety Assessment is consistent with Debra, Keisha, and Ann's descriptions of the role that the risk assessment plays in decision-making. State A's policies create little room for other formal means for making decisions and shape how many workers come to understand the kind of evidence they need to evaluate to form a decision about a case.

More Exclusive Than Mutual: The Role of Clinical Judgment within Actuarial-Based Decision-Making

The Children Research Center's (the developers of the SDM), assert that the SDM should be integrated with investigators' and caseworkers' practice wisdom and

clinical judgment to reach a final conclusion about the disposition of case. The Children Research Center states that:

> By design, SDM assessments do not MAKE decisions (their emphasis). The recommended result of an SDM assessment should be tested against family perspective and worker clinical judgment. When all are aligned, the direction is clear. When there are differences, the SDM model provides a framework for dialogue until there is resolution ("The Structured Decision Making Model: An Evidence-based Approach to Human Services," 2008)

However, State A's legislative requirements to base child welfare decisions off the results of the SDM create a system where both family perspectives and workers' clinical judgments are marginalized in the decision-making process. While the SDM model is technically supposed to offer a platform for integrating clinical case judgment with the results of the actuarial based assessments, in practice, the formal linkage of the risk assessment with the case categories makes it very difficult to leave a space for clinical judgment in the formal decision-making process. When workers disagree with the score on the risk assessment, they often find that policy does not allow a space to incorporate this assessment into the final decision. Angela describes just how little room there is for balancing clinical judgment against the technical judgment of the SDM:

> Sometimes you just feel bad. That's all…you just feel bad and you try to look…like you try to go over the risk assessment again and see maybe if you did something wrong if you clicked on the wrong button. And it's like I don't…'cause it was a case that I opened for a dad and the mom is the one who had this previous history and you know he just married this woman and he did give the kid a whipping and they got a bruise; or something like that. And I felt like, "Oh well he'll be a Category 3 since it's just him. But even though I did a separate risk assessment for him, they still made him a Category 2 because the mom had all this previous history. And so he…[ended up on the Central Registry]

Notice that Angela redoes the risk assessment to see if she "clicked on the wrong button" not because the risk assessment might be incorrect itself. In this way, policy can make the risk assessment appear to be infallible. While workers are permitted to

raise the risk level in cases where they believe this action to be warranted, they are prohibited from lowering it. So even though Angela believes that the father in the case does not represent a high risk of being a repeat offender of maltreatment, she knows that there is no official avenue to reflect this interpretation of the facts of the case. Angela tries to alter the outcome within the rules (by filling out multiple risk assessments) but is ultimately unable to change the final disposition of the case. Casey describes a similar situation for a case that she would like to simply offer services to but where the risk assessment prompts a removal of the children from the home:

> And when, you know I looked at the case because of mom's mental health issues, because she has three children and one of, her youngest baby was special needs. She was in a domestic violence victim, she had CPS history as a victim; all of these things made her have a high risk and made us have to go to the point where we had to file in-home. But she was completely willing to work with services. She was so happy to have someone to help her get out of the domestic violence situation. To help her get her own housing because she's always relied on her mom or a man and this was her first opportunity to support herself and we were helping her and she went to counseling and you know, she went with Families First and this was all before we filed. And just because her risk assessment was so high; we had to file.

Casey goes on to say that she plans to ask for the court to dismiss this particular case but is frustrated by a system that makes it impossible to incorporate the nuances of the information captured on the risk assessment into the final decision. While the risk assessment accurately captures the presence of domestic violence and mental health issues in the family, it does not account for the ways in which the client is working to address these areas of concern that might mitigate the risk of maltreatment reoccurring. Casey summarizes her concerns about the Structured Decision-Making model as it is practiced in State A stating that "…we lose discretion, I feel like, by this assessment program." The formal policy in State A

creates an implementation of the Structured Decision-Making model that does not allow for the integration of clinical judgment alongside it.

The supervision process in State A reinforces the critical role that the SDM has in determining the outcome of a case. While the supervision process might be the ideal space in which to resolve the kind of conflict that Angela experiences, some workers described this space as a place where the authority of the SDM was reified. I asked Dorothy directly about what happened in a case where she opposed the outcome of the SDM. Dorothy described approaching her supervisor to review her concerns:

> What happened with that one is I had asked my supervisor, I said, I don't…"My initial thought was, after doing it [was] I don't think the risk assessment came out right. I don't think this case should be opened. I did not want to open it because I didn't think that mom needed that much help. But when we went over it together and she's like, no, you have to rely on your system…
>
> But she went through and she's like, no, you answered the questions very appropriately and the SWSS [the computer system that administers the SDM] did it right. It needs to open because of all the risk factors there and because mom is saying she's overwhelmed. She's telling you she's overwhelmed so something's there. You need to open it and help the family out.

In this case, the Supervisor's message to Dorothy is that the risk assessment has more merit than her clinical judgment. This kind of exchange reinforces the idea that a worker's own discretion is of little value and that decision-making in State A means being able to accurately document and implement case decisions according to the policy.

However, it is important to note in this example that Dorothy and her Supervisor conflate the outcome on the risk assessment with the requirement to open or substantiate the case. While the risk assessment is meant to determine the case category based on the potential for future maltreatment, it is not supposed to be

used to establish the conditions for substantiation as it is used here. The sense that the risk assessment is the most accurate tool for understanding a case not only overrides this worker's judgment but also overrides the official policy itself, redefining for everyone (not just the two people described here) what kinds of decisions have room for clinical discretion within the context of a structured assessment. In this way, a highly structured official policy for decision-making leads to an informal culture that may make decision-making even more standardized than the original tool intends.

Out of Sight, Out of Mind:

In interviews with participants in State A, the main focus of discussions related to the SDM were descriptions of and reflections on the meaning and use of the Safety and risk assessments for case trajectories. However, the SDM contains multiple components including a Family and Strengths Needs Assessment (FSNA) as well as a Childs Strengths and Needs Assessment (CSNA) for cases that have been substantiated. The FSNA structures the service plan for open cases. Following a similar format to the risk assessment, The Family Strengths and Needs Assessment scores items on a continuum related to family and parental challenges such as substance use, parenting skills, social or community support, and household relationships. The issues that receive the highest number of points are identified as 'priority needs' and determine which services should form the bulk of the service plan. In State A and State B, service plans are required to address the three highest scoring categories. The Child Strengths and Needs Assessment (CSNA) follows the same model as FSNA. As indicated by the name, this assessment focuses on identifying the strengths and needs of the children in the household with the goal of

ensuring that no glaring issues related to the children are left unattended such as developmental delays, untreated medical conditions, or mental health issues. Despite their required use, the majority of participants in State A did not discuss these standardized forms unless prompted to do so. When asked about how they were used in decision-making, workers expressed a range of opinions from "redundant" or "repetitive" to extremely helpful. Keisha describes how theses components of the SDM can be useful in the decision-making process:

> This is our key of what we're actually helping them. This is our services. That's how I look at it. Because whatever the issue is on here, 'cause it's gonna tell you; it's gonna pop up; either housing, substance abuse or you know parenting or something like that. It's gonna show up in plain so you know what you have to deal with. And that helps you when you do the FSNA and the CSNA. You also do the services after that 'cause you know where you gonna go and how you're gonna service them.

While these tools have formal decision-making power in the sense that they structure which serves are prioritized or required for families over others, workers generally did not discuss these tools using the same intensity that they reserved for the Safety and risk assessment (whatever the valence of their assessment to the SDM's utility). For substantiated cases, the Safety and risk assessment holds formal decision-making power over what happens but not formal power over whether something will happen (or not). In this way, use of these tools contain much lower stakes for both the child welfare workers and the families involved in their use. As such, it is likely that they warranted less notice and take up less emotional real estate in discussions about participants' work practices. It makes sense that the tools that hold the most weight for decision-making are the ones that are most likely tone engaged with by workers at a conceptual, practical, and emotional level.

Exceptions to the Rule

It is important to note that not all workers followed the official policy for decision-making in State A. In Same Difference: Standardized Decision-Making in Practice (Bosk, In Progress), I demonstrate that a subset of State A workers (approximately one-third) reported scoring the SDM tools so they captured a decision that they had already made or changing their answers on the SDM to align with their case judgment. Workers who had a more flexible approach to the SDM and who felt entitled to deviate from the SDM tended to have more social privilege in terms of some combination of years or experience, gender, and race. Understandings of the SDM as an effective tool for child welfare decision-making were not necessarily correlated with using the SDM as prescribed. While the majority of workers interviewed used the SDM as intended (approximately two-thirds), only one-third of the workers felt positively about it as a tool for decision-making. Those who felt negatively about the tool tended felt that it was a blunt instrument that did not allow for the integration of tacit and clinical judgment. For a more detailed discussion of these issues, please see Bosk, Same Difference: Standardized Decision-Making In Practice, In Progress).

The SDM in State B: An Administrative Imperative Without Teeth

State B Child Welfare Policy

While State B's legislature mandates the use of Evidence-Based Practices in all social services delivered by the state, it does not currently mandate the specific use of the Structured-Decision-Making Model in its child welfare cases. Additionally, State B has several approaches to assessing and servicing families

referred to the child welfare system. The Differential Response System (DRS) creates a different track for families that appear to be at lower risk for future child maltreatment or where the maltreatment being investigating is not considered to severe. The goal of the differential response system is to engage families in a process that is intended to be supportive and strengths-based. Child welfare policy in State B requires investigators to complete the risk assessment from the Structured Decision-Making model to form an opinion about the disposition of the case. However, unlike State A, this policy does not explicitly require investigators to utilize the score on the risk assessment to determine the trajectory of a case once it has been substantiated. Rather, child welfare workers in State B are asked to evaluate multiple sources of information when making their decision in consultation with their supervisor. The official steps conducted prior to the conclusion of the investigation as described in the Handbook for child welfare investigators are as follows:

"The investigator shall take the steps below to follow-up the field response and conclude the investigation:

- document information gathered during the investigation
- **complete a risk assessment**
- submit a request for a State Police criminal record check on the alleged perpetrator, adults in the home and other adults who have child caring responsibilities or access to the child(ren) in the home, if not done previously
- refer family members for a drug and/or alcohol screen, if needed
- contact collaterals
- **discuss the case with the supervisor, taking into consideration all facts that were collected during the investigation, to determine the plan of action**
- **determine with the supervisor if child abuse or neglect is substantiated or unsubstantiated**
- **determine appropriate Juvenile Court action (see below – Use of Juvenile Court)**
- **take removal actions, when indicated**

- **determine case disposition**
- provide notification of the results of the investigation to those specified below
- utilize the "Investigation Checklist" (DCF-2075) as a guide to ensure that all required actions have been completed."

(State B Department of Children and Families Handbook)

These steps characterize a much looser context for using the Structured Decision-Making Model to inform child welfare decision-making. Workers are required to complete the risk assessment but the policy does not make explicit how the results of the risk assessment should factor into workers case judgments. As the official investigative process outlines, multiple steps (represented in bold) in the decision-making process leave room for caseworker discretion as well as supervisory consultation. In this way, the SDM is implemented within an adaptive framework. This adaptive framework allows for a flexible application of the Structured Decision-Making model that explicitly seeks to integrate the SDM into local culture, office policy, and individual worker and supervisory judgments. However, as the data presented below demonstrates, the flexibility of this process combined with workers negative views about the SDM render it an all but meaningless part of the decision-making process.

An Administrative Imperative Without Teeth

A majority of the 31 respondents in State B reflected a fairly uniform response to the SDM: that it had little to no impact on case decision-making. In initial descriptions of the procedure for investigating or working (in the case of on-going treatment or adolescent treatment workers) a neglect case, the majority of workers did not recount including the SDM as part of their routine. While completing the SDM certainly is a requirement for any case, its lack of salience as

part of the process was an early indication of its relatively limited role in decision-making.

Following the same interview protocol as in State A, I asked participants to walk me through the general process for investigating or working a neglect case. Answering this question, participants were more likely to report aspects of their practice other than the SDM such as: following the priority code (when they needed to make contact with a family), recording their interactions with families and other providers or people relevant to the case, documenting their work through the drafting of case narratives and case reports, and recounting their process for investigating and visiting families, which often included detail about decision-making (e.g. When I visit the house, I make sure the child is safe by "laying eyes on the child" or "I interview the child to find out what is going on").

Participants did include administrative requirements in their descriptions of their work process but the SDM, even as a clerical duty, did not warrant a mention as a central part of their efforts. These descriptions of other administrative duties signaled that participants' conceptualization of the process for working a case included the elements of their job governed by paperwork. Worker's omissions of the SDM are therefore unlikely to be attributable to a desire to hide the less glamorous details of their jobs. Nor are they likely to be indicative of worker's desires to present themselves as always in charge of the direction of a case (Young, 2004; Zerubavel, 2006). The silences around use of the SDM as a decision-making tool suggests that the SDM is not a significant part of the process of a case but functions more as an afterthought, an obligation folded into many other administrative tasks. As such, the SDM is unlikely to be a critical factor in decision-

making for State B's frontline workers.

While omissions about the SDM in work descriptions signaled its lack of decision-making power, direct questioning about the use of SDM in practice tended to confirm this view. When workers were asked how the SDM influenced their decision-making, many articulated that it did not. Consider the following exchange:

> E: How much do they [the SDM tools] drive decision-making?
>
> F: They say that, (short laugh) it…(trails off, long pause)
>
> E: I'm looking for the real, not what they say. But yeah what do they say and
>
> then what happens?
>
> F: This is all confidential right?
>
> E: It is totally confidential.
>
> F: I know, I know, I'm doing the same thing now with getting my Master's and everything, but I don't see it. I really don't think they really drive a case or the determining factors to you know to move the case along for closure or stay involved. Yeah. Are we closing the case, no, no, not at all. I can probably say on half of my caseload right now, like I said the risk level is low and we're staying involved.

Felicia, an African-American investigator identifies with the perspective of the researcher through her status as an MSW student. She understands the theoretical purpose of standardization, noting, "she is doing the same thing now with getting my Master's" but in practice does not observe that the actual benefit of the SDM tools lines up with their promise. Felicia assumes that researchers, like myself have a positive view of the SDM. Despite her implicit acknowledgment that the 'right answer' is that the SDM is effective, Felicia reports that the SDM does not act as intended and is not part of the decision-making process. Seeking reassurance that the content of the interview will not be passed on to her superiors, she acknowledges that the informal work practices related to the risk assessment do not align either with formal departmental policy or with the explicit justification for

caseworkers to utilize the tools. The informal practice may predominant caseworker actions but the official narrative of what the SDM achieves remains an important part of its presentation. Christiana an on-going treatment worker echoes this perspective:

> Oh when this tool first came out it was more so like yes, if you have a case that is low from, I don't know, two consecutive times or whatever the case might be, then you should just close, but now I think that has kind of like been like slipped under the rug or somewhere.

Christiana reports that the way the RA is currently handled in every day practice conflicts with its intended use. The SDM itself has not been slipped under the rug, it is a key requirement of the work for every case, what has disappeared out of plain view are the ways that the SDM could or would impact case trajectories.

The Role of Supervision

While official policy may dictate use of the SDM in guiding case decisions, a critical part of its implementation is the weight that the SDM carries in dialogues between supervisor and caseworkers. If the SDM held much sway in the conversation about case outcomes, then we would expect that the SDM would be a regular part of the supervisory process. Both caseworkers and supervisors reported that the SDM was more likely to play a confirmatory role in case decisions. Melanie an on-going treatment worker explains that supervision rather than the SDM offers her the direction for a case.

> M: Would you ever use this to help you make a decision about what you should do in your case?
>
> P: To some degree.
>
> M: Can you give me an example of a time where you've used it?
>
> P: Hmm. I can't think of the last time. Because really, I mean you're not really doing that risk assessment until you're like finalizing your investigation. And most of the time you've already at least talked to your supervisor at least once…

M: Okay.

P: …in supervision, about the direction you think the case is going.

Melanie's account embodies the disconnection between official language about the purpose and usefulness of the SDM and the effect that the SDM has in practice. She begins by reporting that the SDM is "somewhat" useful in decision-making but can't remember a time that the SDM steered her towards a particular outcome. Unable to come up with a case example, Melanie switches her account to an explanation of how the tools actually play a relatively minor role in decision-making. Discussions with Melanie's supervisor give her an initial idea for how the case should (and will) proceed. Beth concurs:

B: I mean, I guess it would be an interesting question to ask if supervisors
actually looked through it or if they just approve it.

E: Something, I'm asking.

B: Yeah, because I've never had a supervisor come to me and say; I don't agree with this. Not once has that ever happened. So it just gets filtered as like part of the documentation….

Because meetings where case trajectories get hammered out may not or do not include the SDM as resource or reference point, there is little reason for workers who do not discuss it as part of the case review to think that the outcome on any of the SDM tools matter for casework. The workflow patterns in State B offices indicate that tools requiring completion at the conclusion of a work process may mean that workers unintentionally develop work practices that exclude them from decision-making.

According to the majority of workers, the SDM not only plays a relatively minor role in routine decision-making about cases, it also is not widely utilized to help figure out how to proceed in more complex situations. Lisa, like Melanie, sees supervision as the more appropriate venue to decide how to handle complicated

issues.

> M: When you are on the fence about something what do you do? What do you use to figure out what you've ---?
>
> L: When I'm on the fence, me I kind of go with what I feel, with my gut and you know, I'll throw it at my supervisor. Like I said he's very receptive to what I need and what I need from him, and if I need to sort of throw something off him you know cause I'll say I really go to him if I need to be told to do something. It's usually I kind of have a good idea, I put it all together, I'll sometimes even write it up without us even conferencing it, and he'll say if I have a question I'll come see you about it, but for the most part he's comfortable with it and I'm comfortable with it. I'll go to him when I'm kind of stuck or a little unsure about something and he'll push me along kind of.

One clear rationale for using standardized tools is to eliminate or at least balance decision-making that comes from a strong emotional perspective. Lisa does not use the SDM as a double check for her feelings even though there is the opportunity to do so. Nor is Lisa solely reliant on "her gut" to determine what should be done, she makes sure to present her thinking about the case to her supervisor and uses this guidance to get unstuck. Curious about whether Lisa's supervisor made use of the SDM as a reference point in either these exchanges or more routine conversations about case decision-making, I asked whether her supervisor incorporated review of the SDM tools into their case conferences. Lisa responded: "I mean to some degree he does use it. I've never had him kick a case back to me saying 'Lisa, the risk assessment is moderate why are we doing this you know?'.

While discussions about the SDM can certainly be a part of larger conversations between supervisors and workers about the direction of a case, the tool's informal lack of power becomes underscored when there is little disagreement or larger reflection about various scores on SDM assessments in supervision. Accepted uncritically, worker's scores on the assessment become folded in with the

large volume of paperwork that accompanies any case. While there might be large back and forth discussions about case trajectories in conversation, perfunctory engagement with the SDM between supervisors and investigators signals its relatively low status for offering an important viewpoint on cases.

Workers aren't the only ones who feel that the SDM functions in a perfunctory fashion. Supervisors I interviewed tended to echo this understanding. Agnes, described the SDM in strikingly similar terms to Felicia and Christiana. Asking directly about the SDM led to the following exchange:

> E: What role does that tool [the Risk Reassessment and Reunification Assessment] specifically play in your decision-making?
>
> A: Well for me right now…the reality is it's supposed to be that you're looking at it and you're really using it to help whether you're gonna reunify someone. So the reality is it plays a small role...
>
> [A follow-up question was asked later in the conversation]
>
> E: ….Sure. I just wanted to clarify: By the time the worker gets to filling out the reunification…and the reassessment form, has the decision already been made?
>
> A: Most of the time…
>
> E …And then this [the Reassessment and Reunification tools] is sort of like a confirmation?
>
> A: Exactly.
>
> E: Okay.
>
> A: More like a confirmation. And sometimes, I feel it's like just being done because it's due.

Like the workers mentioned above, Agnes recognizes that there is a difference between how the SDM should be used (according to formal policy) and how it operates in the everyday routines of workers and their supervisors. Acting as "confirmation," the SDM can provide insight into whether a worker's and supervisor's decision-making align with more standardized assessments. However, Agnes's feeling that these assessments are often only used to meet a deadline somewhat undercuts this purpose. Agnes explained that time pressures interfere

with her ability to really review the SDM tools to inform decision-making:

> Right, right. 'Cause if we were able to actually look at what we're documenting; like what we're supposed to supervise; everything that's documented on a case we're supposed to look at. If we actually had the luxury of looking at all that it might change the direction a case is going; it really might. It might change it for the better, it might change it for the worse.

For Agnes, the time and space to actually review every piece of documentation associated with a case is an unrealistic 'luxury.' Choosing to prioritize her time reviewing cases with workers orally, Agnes understands the SDM tools to be an extra step in supervision that is out of range for what is manageable on any given day. Agnostic about whether the SDM improves or impedes good decision-making, Agnes simply attributes their lack of use to lack of time. How workers understand tools to fit into their overall workplace demands may be a critical variable for understanding how structural and organizational factors impact their overall use. No matter how well a tool is constructed, if it does not fit in with an overall work plan or flow its uptake is likely to be limited.

Other supervisors also privileged conversational and reflective processes to inform decision-making. Ellen believes that discussions with her supervisees provide the most effective sources of information about a case:

> I mean I feel like, you know I do a lot of very, of supervision a lot with my workers, so we're going through the case for the day and I only meet with them for 10, 15, 30, you know and I don't do what other supervisors do because I want to be part of it and I'm owning the case too. So I'm kind of feeling like where they're at and I can help them figure out another way to go or something like that so I don't feel like I need the paperwork.

Its not the paperwork (and she includes the SDM in this category) that makes Ellen feel that she "owns" a case with a worker but her attention to the story of what is happening within it. This is an interesting distinction to make as one of Ellen's primary tasks as a supervisor is to ensure that the documentation for a case has been

completed and turned in on time. Such a distinction indicates that some supervisors may believe that tools like the SDM do little to add value to decision-making. Here, Ellen performs both her role as an expert and her authority by using clinical judgment to guide the investigator.

In addition to discussions with their supervisors, workers in State B had access to additional methods for assessing cases. Teams composed of medical, legal, and psychological experts in specific areas such as trauma and domestic violence were readily available to consult with workers on how these issues might be impacting their cases. Workers initiated these consultations when they felt that their case required additional thinking or when it was complicated to understand how to proceed. Participants reported placing great weight on the information they received from these discussions. Further, a team of physicians, social workers, and psychologists regularly review the physical abuse cases appearing in the system. The end result is an organizational approach that structures regular opportunities for evaluating risk in multiple ways from the risk assessment itself to expert discussions. When there are other structural opportunities for thinking about the meaning of risk in a case, the SDM may become devalued.

A Bureaucratic Tool For Management

The process by which workers describe utilizing the SDM in their casework is indicative of the extent to which caseworkers in State B engage with the SDM as a bureaucratic tool. As discussed above, in the majority of interviews workers described filling out the structured decision-making assessments after a decision had been made as part of the paperwork required to finalize the opening, closing or transfer of a case. Workers reported experiencing the SDM as a part of the every

day paperwork required by her job. Pat recounts that:

> The SDM that at one point it was something that was coming down in LINK [the computer system in State B that tracks cases] and we were going to be expected to do it and then it was mandatory because now LINK won't let you sign off on a case unless you do it; it'll jump you back.

Presented as a rote task, the SDM is necessary only in that one has to be filled out before being able to simply submit work via the computer system. When asked directly about the SDM, workers wondered aloud about its purpose. Pat finished her thoughts about the SDM with: " So this tool, I don't know what, this is just something I think managers need or something or agency; I don't know who needs it." Workers often were unclear on how the SDM fit in with the larger picture. Some workers hypothesized that the SDM's function was really not for workers but for managers who needed to gather aggregate level information of cases. As Grace, a noted:

> With the SDM tools, I really honestly, I think it, those tools are really for numbers and management. It's not something that helps us determine, us as workers, determine what needs to happen on a case.

As an administrative imperative, SDM was not seen as being useful or important to individuals' own work or decision-making processes. Instead, the SDM was conceptualized as something that management "did," clearly demarcated from what frontline workers "do." Investigators and on-going treatment workers in State B described a rigidly hierarchical environment where job title determined status, pay, duties performed on the job, and formal decision-making power. The sense that SDM was for management rather than for workers contributed to CPS workers' conceptualization of SDM tools as not useful or a "real" part of their everyday work. SDM became categorized as solely an administrative function to reflect this division of labor.

Descriptions about SDM being "about the numbers" were folded into much broader complaints about the direction of caseworker experience in these two offices in general. Workers placed their critiques about the SDM within a larger context of casework driven by bureaucratic imperatives and measurements. Mike, an investigator, perhaps put it the most forcefully, but many other workers echoed his sentiments.

> M: That's all it is, it's just paperwork after paperwork after paperwork, and I feel like sometimes my visits with families suffer, because you need to be in, be out, check you know the safety factors, make sure that the kids are, you know got food in the home, all utilities working and check the kid up and down to see if there's any marks, bruises, depending on what the case is. And then it's bang, bang, bang, bang, I'm out and that's why I think a lot of clients are little bit resistant at first. I try to spend more time with them, but then my written work suffers and here in this particular office I think it's all about numbers and it's all about you know getting our benchmarks on everything is kind of like, but obviously what the state wants, but one thing more that the managers care about, even though they say otherwise, I call BS on it.
>
> E: They just want the numbers?
>
> M: It's all about the numbers.
>
> E: Which numbers? What are the ones that are really important to them?
>
> M: It's the visitations, you know we have to achieve 85% of the benchmarks every month. Yeah, I don't have any problem doing that, but God forbid you miss one you run the risk of getting written up.

Mike makes the issues related to the use of the SDM wider than just these tools. Distinguishing what is important to management (numbers, benchmarks) and what is important to him (building a relationship with the family, spending time with them), Mike positions more quantitative measures as something different from the work that he believes is meaningful and reflective of his professional identity. Struggling with the competing demands imposed by two very different requirements of her job, Mike frames them in either/or terms. It is quite possible

that worker perspectives on the SDM become folded into a larger contentious climate about administrative paperwork. Workers negative feelings about the bureaucratic imperatives they are required to follow may not be easily separated from their views on the SDM. These feelings are then overlaid by what they experience to be an unforgiving administrative environment. Mike indicates how punitively he feels the administrative parts of his job are and is aware that if he is late or misses anything that he "run[s] the risk of being written up" for not doing his job correctly. Penalizations for not completing written work further contribute to a sense that some types of work are for supervisors and managers and others are for workers rather than a shared experience toward a common goal.

Agnes, a supervisor commented on how difficult such an environment can be:

> Sure. People are very, very stressed by…I mean numbers, you have to make sure of your visits. You have to make sure you're visiting twice a month or you have to have a certain percentage of your narratives in. And so the workers will feel as though, "Well I've got all this work to do. How can I possibly keep up?..." So the numbers game is really stressful for people and you feel it.

In State B, the paperwork, numbers, and stress are tightly connected to one another. As Mike put it later in the interview: "They think you're actually doing social work and I really, I think we do, we don't do much of it. In response to my question "What do you do?" he answered: Paperwork." Many workers like Mike, Grace, and Pat envision their real work as the tasks that call on their clinical judgment, professional expertise, and understanding. Paperwork is a necessary evil of the job, an evil that calls up a certain amount of resistance in order to hold onto a distinct feeling of professional authority.

A Paper Zoo

As part of my fieldwork, I had the opportunity to observe workers making case decisions over a month period. Workers had folders filled with assessment instruments that they were encouraged to use when evaluating cases such as a specific inventory for the presence of domestic violence, another inventory for detecting depression, and other tools that had come in and out of fashion at different moments in state or office policy. In the sea of paper, the cumulative effect was to render each tool meaningless and less likely to be used. As State B policy moved towards engaging families from a strengths-based perspective, even more tools were being added to partner with families during the child welfare investigation and treatment process. Whenever a new intervention was introduced to better facilitate the process, old interventions were not taken away. Workers described being overwhelmed by the number of interventions intended to facilitate better decision-making and the sense I had was that they were surrounded by "paper tigers." (Timmermans & Epstein, 2010).

It is important to note and reflect on the fact that some workers and supervisors *did* feel that the SDM was a useful tool for guiding case actions and utilized it as such. Marie, who felt overwhelmed by the paperwork aspects of her job reported that she still used the SDM tools to inform case actions: "Yes. It gives you the direction of a case and it helps you process, you know what the needs are, what the strengths are and in the direction of where it needs to go." For Marie, the SDM assists in understanding the full range of issues in a case, which in turn facilitate critical thinking about how to proceed. Zana shared this perspective:

M: And what is particularly helpful about the risk reassessment?

P: Well it helps you to determine if the case could be closed or not. It give you the level of risk, maybe low, moderate or so on, and of course if it's moderate there is something that is missing, something that they need to address; maybe the parents are refusing, which there's nothing we can do or they still need to continue whatever services. And it really determines what we should do with the case.

Zana values the set of questions that the SDM asks as a way to organize her case thinking and she links this case thinking to case outcomes.

Exceptions to the Rule, Here Too

Marie and Zana represent the small number of workers I interviewed who utilized the SDM as intended. Their accounts of the SDM are positive and engaged. For these workers, the SDM functions as a source of valuable information rather than as a redundant clerical task. I also observed a Program Manager direct two investigators to use the risk assessment to determine whether the case should be substantiated and opened when they came to her for help in making a decision in a complex case. Natasha also described using the SDM with her workers when she felt that their decisions were based in emotion rather than the facts of the case. She also regularly insisted that her workers use the SDM for determining case actions when decisions were unclear or there were disagreements about how to proceed. However, these approaches were very much in the minority in how the SDM was routinely applied to casework. As such, Marie, Zana and Natasha's descriptions of the usefulness of the SDM exemplify individual variation in workers' perspectives but do not provide insight into the normative culture surrounding these tools. Overall, both offices in State B developed an informal practice of utilizing the SDM for confirmation of an already made decision or folded it into administrative tasks that were required but were largely un-reflected and un-acted upon.

One of These Things is Not Like the Other: The Local Context of Standardization

In this article, I have demonstrated that the Structured Decision-Making model is sensitive to organizational contexts. As a practice technology, the significance of the SDM becomes socially created in every setting in which it is deployed. Rather than standardizing work practices across all child welfare offices, how workers understand and then utilize the SDM reflects a variety of organizational factors and orientations. Among these factors are: (1) the weight given to the SDM in guiding specific outcomes for cases (e.g. opening or closing a case, level of substantiation or whether the alleged abuser is placed on the Central Registry for Abuse and Neglect);[1] (2) whether supervisors review caseworkers decisions in conjunction with the findings of the SDM; (3) how workers view the administrative response to the SDM; and (4) the informal culture for completing administrative work and making decisions in a protective services office.

State policy structures how the SDM will be used in practice. The distinct constructions of the SDM by workers in State A and State B illustrate the intensely local nature of tools designed to be universal in their application. Investigators and treatment workers in State B report that the SDM plays a relatively minor role in their decision-making overall. On the whole, workers in this state experience the SDM as an administrative procedure utilized by management for its ability to capture meta-data and aggregate trends of cases. The SDM represents one set of procedures to make the audit of performance possible (Power, 1997). While workers are required to complete each part of the SDM, the results do not have to be tightly correlated with the disposition of the case. In State B, the SDM functions more as a

suggestion for the direction of the case, weakening the connection between the meaningfulness of the information it provides and the decision-making the SDM is intended to direct. The looser the formal link, the less likely workers are to view the SDM as anything other than another part of their required documentation. A minority of workers and supervisors felt the SDM had tremendous value and used the tools to guide their decision-making. However, workers and supervisors who relied on the SDM to inform case thinking represented individual variation in approaches to casework rather than institutional norms to incorporate the SDM into decision-making processes. The presence of other avenues for case decision-making that involved consultation with experts as well as the sheer number of tools meant to facilitate accurate and empirically-based decision-making may decrease the importance of the SDM in practice.

In contrast, the SDM in State A has been deeply institutionalized, shaping case outcomes and case trajectories. Scores on the risk assessment determine the intensity of services families will receive, the category at which a case is substantiated, and whether a family is placed on the state's Central Registry for Abuse and Neglect. In some instances, the score on a risk assessment can even contribute to the decision to open a case that falls on the border for meeting the criteria for substantiation. Workers uniformly acknowledge the significance of the SDM for decision-making but vary in how they approach the tool. No matter what their personal orientation to the SDM, every worker I interviewed recognized and remarked on the formal power the SDM had to determine case trajectories.

Differences between the implementation of the SDM in these states demonstrate that whether a standardized tool is understood to achieve a

standardizing function is not a consequence of the validity or reliability of the tool itself but the social and structural context in which a tool is introduced and utilized (Dixon-Woods, et al., 2011; Epstein, 2007; Espeland & Stevens, 1998; Espeland & Vannebo, 2007; Timmermans & Epstein, 2010).

Often viewed as a natural, mechanistic process, the data presented here challenges the idea that standardization is anything but a social practice whose meaning is shaped by factors external to the tools themselves. A highly contingent phenomenon, the implementation of the SDM in State A and State B indicates that responses to standardization are shaped through a dialectical process. The introduction of formal policy related to the SDM sets the expectations for a set of work practices that are then renegotiated by workers and supervisors in their every day work routines. In turn, workers every day approaches to *formal* policy create an *informal* practice that guides the use of standardized tools. Once an informal approach has been established, this way of doing things rather than the formal policy becomes the normative and expected practice. The meaning and use of the SDM become 'the way we do our work here.'

This meaning and use is, as previously stated, highly contingent. Factors such as a change in workplace supervisors, upper-level administration, the addition of other policies, or a service 'failure' that becomes a public scandal—all have the possibility of changing workplace practices related to the SDM.

Viewing standardization as a social process is critical to understanding the creation and implementation of interventions that are conceptualized as operating similarly across settings. Recognizing that standardized tools function differently in distinct organizational contexts raises questions about whether standardization

itself will erase or even mitigate the messiness of complex work. Further, in the child welfare literature, disparities in child welfare outcomes are often investigated and then addressed at the level of the individual worker. However, my research shows that variation in individual worker's decision-making occurs within larger variation in work practices at the organizational level; variation that itself is nested within differences in policy across states and the federal government. In the cases that I present above, state policy changes the meaning of the SDM in each setting, which in turn shapes the meaning and use of the tool by individual workers. This article unpacks how variation in policy at the organizational level changes the meaning of the SDM on the ground in order to understand standardization as adaptive rather than technical work.

The Goldilocks Problem: Adaptive vs. Restricted Implementation

Implementation Science has emerged as a field to respond to the recognition that interventions with great internal validity will not necessarily translate into external validity in the settings in which they are deployed. Across the fields of medicine, psychology, and social work, research has drawn attention to the ways in which EBPs have gaps between the intervention as designed and the intervention as it is practiced (Westen, Novotny, and Thompson-Brenner 2004, 2005; Magill 2006; Howard et al. 2009; Yunong and Fengzhi 2009; Berger 2010). Implementation science seeks to identify the factors that will promote the successful deployment and sustainment of interventions. A major concern of the field is to develop systematic knowledge and strategies to address known challenges to translating EBPs into diverse practice settings. A number of studies investigating the implementation of Evidence-Based Practice interventions in medicine, psychology, and social work

have demonstrated that simply providing intensive training and education are necessary but not sufficient conditions for the successful diffusion of an EBP. (Beidas & Kendall, 2010; Herschell, Kolko, Baumann, & Davis, 2010). Research suggests that for implementation efforts to succeed multiple variables need to be attended to in addition to adequate training and support such as the organizational culture and context, appropriate infrastructure and technological capacity, and leadership.

Implementation science as a field assumes that with the right combination of variables and with careful enough attention to environmental, organizational, and individual factors then interventions will work as intended. The trick (and therefore the research agenda) is to uncover what those adjustments and variables need to be. Recognizing the importance of incorporating the organizational and cultural context in the implementation of interventions raises questions about whether an adaptive vs. restricted implementation strategy will be the most effective.

However, it is not clear that such a task is even possible. As this article demonstrates the implementation and organizational context will shape how an intervention is deployed in ways that shape the goals of or change the meaning of the intervention itself. Reviewing sociological work on standardization: Timmermans and Epstein state:

> As Wittgenstein (1953) and ethnomethodologists have noted (Heritage 1984), no rule can adequately capture the requisite work of a prescribed action. On the ground, every standard is simultaneously over determined and incomplete. To coordinate diverse interests and activities, standards necessarily delegate some residual work that requires active participation and submission of people to the standard's directives. Tinkering, repairing, subverting, or circumventing prescriptions of the standard are necessary to make standards work (Lampland & Star 2009, p. 4; Star 1995, pp. 100–104). Thus, a recurring surprising finding is that loose standards with great

> adaptability may work better than rigidly defined standards (Timmermans & Epstein, 2010).

An open question is what does "a loose standard with great adaptability" looks like in practice. State B employs a loose standard with great adaptability by requiring the SDM to be part of the assessment process without formally linking it to case outcomes. This model allows for local culture and context to be taken into account but because workers and supervisors both are overwhelmed by paperwork and do not see value within the SDM itself. The very flexibility that might make the intervention successful ends up transforming the SDM into a bureaucratic exercise with little practical meaning. Timmerman's and Epstein caution that standards cannot be so flexible that they lose all their meaning (Timmermans & Epstein, 2010).

Taken in isolation, analyzing the implementation of the SDM in State B and State A might lead to opposite conclusions. Theorizing about how to make the SDM a more utilized intervention in State B, one might suggest tying use of the SDM to case trajectories. In contrast examining the data in State A might lead to suggestions for policy improvement that would decouple the tight relationship between the score on the SDM and the trajectory of the case. However, when looking at both approaches together, it is unclear that either works as well as it should. State B's implementation approach is so adaptive that workers can choose to ignore the intervention completely, while State A's approach is so restrictive that it is difficult to allow clinical judgment into the process. I have come to think of this as the Goldilocks problem. What is the right amount of adaptability vs. fidelity to the model that should inherently be part of the implementation process? When examining the list of factors that need to be attended to in successful

implementations is there anything that is not on the table? If so, how is it feasible to offer standardized processes that are realistically adaptable to every day life?

A sociological understanding of standardization as both a social and technical process requires close attention to context for the successful deployment of an intervention into real world settings. But what does one do then with the understanding that policy, organizational, and environmental factors will influence the deployment of a new intervention? While there is growing recognition that a balance between flexibility and fidelity to models must be struck for interventions to be successful, implementation science has yet to provide sound or realistic models for what this mix of adaptive vs. restricted should be or how this understanding can be built into interventions. The Practice-Based Evidence movement, which has emerged in response to the messiness of practice in the real world asserts:

> Greater attention to context may require novel uses of research approaches that are closer to the ground and better at capturing the nuances of client preferences and treatment contexts, or combining these methods in novel ways with highly controlled studies. Practice-based evidence, a movement that is developing in psychology, medicine, and education, treats "the multiplicity of variables as a resource to be used for deeper analysis rather than as a nemesis to be controlled" (Green 2006, 406), providing new possibilities for incorporating context into EBP research and knowledge (Green, 2006; Okpych & Yu, 2014).

However, concrete examples or suggestions for what these new methods or novel forms of analysis are have yet to be fleshed out.

In his examination of how state efforts to improve the human condition have failed, James C. Scott, a political scientist and anthropologist offers some insight into how we might conceptualize the relationship between the multiplicity of variables required for an intervention to be translated successfully into the construction of interventions themselves. Scott proposes that knowledge can be thought of in terms

of metis-"practical skills, common sense, experience, know how" and techne-"technical knowledge in the form of hard and fast rules"(Scott, 1998). While metis is often considered suspect in a positivist paradigm, Scott demonstrates that incorporating metis with techne is essential to the success of modernist ideas and plans. Giving examples as varied as the failure and then revival of the planned city of Brasilia to the development of compulsory villages in Tanzania, Scott demonstrates that interventions solely relying on techne are likely to fail because they do not account for particularity, locality, or the complexity built into any set of circumstances. "As Pascal wrote, the great failure of rationalism is 'not its recognition of technical knowledge but its failure to recognize any other." (Scott, 1998, 340). When interventions are all techne they are likely to fail precisely because they do not leave a way to address the unplanned and the unaccounted for. What appears straightforward in one setting or with one case might be difficult to decipher or implement in another. I only need to think to my own experiences following a plan to sleep train my twins to find the truth in this idea. On its face, sleep training required only techne in the form of a clear protocol for responding to night crying in progressive intervals. In practice, this required a great deal of metis as my husband and I tried to work through unplanned events: what should we do when my son peed through the bed? What should we do when his crying was for a short period of time? A standardized process required a certain degree of judgment within the protocol.

As I have argued in the other work (Bosk, Same Difference: Standardized Decision-Making in Practice, In Progress), the Structured Decision-Making model might strike a better balance between metis and techne if there was space to

incorporate narrative reflections about a case alongside the tool itself. A process that invites critical thinking within standardized criteria would create a standard way to think about cases rather than a standardized outcome for cases that might seem similar on their face but are actually quite different in reality. In doing so, workers in State B could not ignore the empirically based risk factors when evaluating their cases nor would workers in State A feel that they had to cast aside nuanced understandings of the dynamic circumstances that influenced the impact of the presence of these risk factors on the potential for further maltreatment.

Creating a set of standards and a standardized process for approaching cases might make more sense then developing a tool around standardizing outcomes when such diversity and complexity exists within the cases themselves. Standardized interventions must account for not only the diversity and complexity of the situations that they are intended to address but also for the diversity and complexity of all the contexts within which they are deployed.

[1] Central Registries of Abuse and Neglect document perpetrators of serious child maltreatment. Maintained at the state level, central registries prohibit those on it from working in jobs that come into contact with children (e.g. school bus driver, day care worker) and in some cases from attending any events at a school (such as a field trip or auction). The decision-making process for who is placed on a central registry varies by state. As of 2008, 40 states had Central Registries

[2] Additionally, State B has several approaches to assessing and servicing families referred to the child welfare system. The Differential Response System (DRS) creates a different track for families that appear to be at lower risk for future child maltreatment or where the maltreatment being investigating is not considered to severe. The goal of the differential response system is to engage families in a process that is intended to be supportive and strengths-based

[3] Bolded items represent my emphasis.

Works Cited

Allan, J. (2006). *Privilege, Power, and Difference*. New York: McGraw-Hill.

Baird, C., & Wagner, D. (2000). The relative validity of actuarial and consensus based risk assessment systems. *Children and Youth Services Review, 22*, 839-871.

Baird, C., Wagner, D., Healy, T., & Johnson, K. (1999). Risk assessment in child protective services: consensus and actuarial model reliability *Child Welfare, 78*(6), 723-748.

Barnett, D., Manly, J., & Cicchetti, D. (1993). Defining child maltreatment: The interface between policy and research. In D. Cicchetti & S. Toth (Eds.), *Child abuse, child development, and social policy*. Norwood, NJ: Ablex.

Charmaz, K. (2006). *Constructing grounded theory: a practical guide through qualitative analysis*. Thousand Oaks: Sage.

Coulton, C., Corbin, J., & Su, M. (1995). Community level factors and child maltreatment rates. *Child Development, 66*, 1262.

Davidson-Arad, B., & Benbenishty, R. (2010). Contribution of child protection workers' attitudes to their risk assessments and intervention recommendations: a study in Israel. *Health & Social Care in the Community, 18*, 1-9.

Dawes, R. M. (1988). *Rational choice in an uncertain world*. Orlando: Harcourt, Brace, Jovanovich.

Dawes, R. M. (1999). A message from psychologists to economists: mere predictability doesn't matter like it should (without a good story appended to it). . *Journal of Economic Behavior & Organization, 39*(1), 29-40.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical vs. actuarial judgment. *Science, 243*, 1668-1674.

DePanfilis, D. (1996). Implementing child mistreatment risk assessment systems: lessons from theory. *Administration in Social Work, 20*(2), 41-59.

Dettlaff, A. J., Rivaux, S. L., Baumann, D. J., Fluke, J. D., Rycraft, J. R., & James, J. (2011). Disentangling substantiation: The influence of race, income, and risk on the substantiation decision in child welfare. *Children and Youth Services Review, 33*(9), 1630-1637. doi: 10.1016/j.childyouth.2011.04.005

Dettlaff, A. J., & Rycraft, J. R. (2008). Deconstructing Disproportionality: Views From Multiple Community Stakeholders. *Child Welfare, 87*(2), 37-58.

Dettlaff, A. J., & Rycraft, J. R. (2010). Factors Contributing to Disproportionality in the Child Welfare System: Views from the Legal Community. *Social work, 55*(3), 213-224.

Dorsey, S., Mustillo, S., Farmer, E., & Elbogen, E. (2008). Caseworker assessments of risk for recurrent maltreatment: Association with case-specific factors and re-reports. *Child Abuse & Neglect, 32*, 377-391.

Dworsky, A., White, C. R., O'Brien, K., Pecora, P., Courtney, M., Kessler, R., . . . Hwang, I. (2010). Racial and ethnic differences in the outcomes of former foster youth. *Children and Youth Services Review, 32*(6), 902-912. doi: 10.1016/j.childyouth.2010.03.001

Emerson, R., Fretz, R., & Shaw, L. (1995). *Writing Ethnographic Fieldnotes*. Chicago: University of Chicago Press.

England, P., McClinktock, E., & Shafer, E. (2011). Birth control use and early, unintended births: Evidence for a class gradient. In M. Carlson & P. England (Eds.), *Social Class and Changing Families in an Unequal America*. Stanford: Stanford University Press.

English, D., & Pecora, P. (1994). Risk assessment as a practice method in child protective services. *Child Welfare, 73*, 451-473.

Espeland, W. (1997). Authority by the numbers: Porter on quantification, discretion, and the legitimation of expertise. *Law and Social Inquiry, 22*.

Evans, T., & Harris, J. (2004). Street Level Bureaucracy, Social Work and the (Exaggerated) Death of Discretion. *British Journal of Social Work, 34*, 871-895.

Fluke, J., Yuan, Y., Hedderson, J., & Curtis, P. (2003). Disproportionate representation of race and ethnicity in child maltreatment: Investigation and victimization. *Children and Youth Services Review, 25*.

Gambrill, E. (1990). *Critical thinking in clinical practice: improving the accuracy of judgments and decisions about clients*. San Francisco: Jossey-Bass.

Gambrill, E., & Shlonsky, A. (2000). risk assessment in context. *Children and Youth Services Review, 22*(11-12), 813-837.

Gambrill, E., & Shlonsky, A. (2001). The need for comprehensive risk management systems in child welfare. *Children and Youth Services Review, 23*(1), 79-107.

Giovannoni, J., & Becerra, R. (1970). Defining Child Abuse. New York: Free Press.

Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical statistical controversy. *Psychology, Public Policy, and Law, 2*(293-323).

Holland, S. (2000). The assessment relationship: interactions between social workers and parents in child protective assessments. *British Journal of Social Work, 30*, 149-163.

Holland, S. (2001). Representing Children in Child Protection Assessments. *Childhood, 8*(3), 322-339.

Hutchinson, E. (1990). Child Maltreatment: Can it be defined? *Child Welfare, 49*(4).

Jones, L. (1993). Decision making in child welfare: A critical review of the literature. *Child and Adolescent Social Work Journal, 10*(3).

Klinenberg, E. (2003). *Heatwave: A social autopsy of disaster in chicago*. Chicago: University of Chicago Press.

Lindsey, D. (1992). Reliability of the foster care placement decision: A review. *Research on Social Work Practice, 2*(1), 65-80.

Loman, A., & Siegel, G. (2004). An Evaluation of the Minnesota SDM Family risk assessment. St. Louis, Missouri: Institute of Applied Research.

Martin, K. (1998). Becoming a Gendered Body: The practices of preschool. *American Sociological Review, 63*(4).

McDonald, T., & Marks, J. (1991). A Review of risk factors assessed in child protective services. *Social Service Review, 65*(1), 112-132.

Munro, E. (1996). Avoidable and unavoidable mistakes in child protection work. *British Journal of Social Work, 26*, 793-808.

Munro, E. (2004). A simpler way to understand the results of risk assessments. *Children and Youth Services Review, 26*, 873-883.

Osmo, R., & Benbenishty, R. (2004). Children at risk: rationales for risk assessments and interventions. *Children and Youth Services Review, 26*, 1155-1173.
Portwood, S. (1998). The impact of individuals' characteristics and experiences on their definitions of child maltreatmen. *Child Abuse & Neglect, 22*(5).
Pugh, A. (2009). *Longing and Belonging: Parents, children and consumer culture.* Berkely: University of California Press.
Rose, S. J. (1999). Reaching Consensus on Child Neglect: African American Mothers and Child Welfare Workers. *Children and Youth Services Review, 21*(6).
Rose, S. J., & Meezan, W. (1996). Variations in perceptions of child neglect. *Child Welfare: Journal of Policy, Practice, and Program, 75*(2), 139-160.
Schwalbe, C. (2004). Re-visioning risk assessment for human service decision making. *Children and Youth Services Review, 26*, 561-576.
The Structured Decision Making Model: An Evidence-based Approach to Human Services. (2008). Madison: Children's Research Center.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainity: hueristics and biases. *Science, 185*, 1124-1131.
Tversky, A., & Kahneman, D. (2002). Extensional vs. intuitive reasoning: the conjunction fallacy in probabilistic judgment. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Hueristics and biases: the psychology of intuitive judgment.* New York: Cambridge University Press.
Wald, M. S., & Woolverton, M. (1990). Risk assessment: The emperor's new clothes? *Child Welfare, 69*(9), 483-511.
Walsh, K., Bridgstock, R., Farrell, A., Rassafiani, M., & Schweitzer, R. (2008). Case, teacher and school characteristics influencing teachers detection and reporting of child physical abuse and neglect: Results from an Australian survey. *Child Abuse & Neglect, 32*(10), 983-993. doi: 10.1016/j.chiabu.2008.03.002
Weiss, R. (1995). *Learning From Strangers: The Art and Method of Qualitative Interview Studies.* New York: The Free Press.
Young, L. (1964). *Wednesday's children: A study of child neglect and abuse.* New York: McGraw-Hill.
Zuvarin, S. J. (1999). Child neglect: A Review of Definistions and Meausrement in Research. In H. Dubowtiz (Ed.), *Neglected children: Research, practice, and policy.* Thousand Oaks: Sage.
Adams, K. B., Matto, H. C., & LeCroy, C. W. (2009). Limitations of Evidence-Based Practice for Social Work Education: Unpacking the Complexity. *Journal of Social Work Education, 45*(2), 165-186.
Ammerman, A., Smith, T. W., & Calancie, L. (2014). Practice-Based Evidence in Public Health: Improving Reach, Relevance, and Results. *Annual Review of Public Health, Vol 35, 35*, 47-63. doi: 10.1146/annurev-publhealth-032013-182458
Atherton, C. R. (1993). Empiricists Versus Social Constructionists - Time for a Cease-Fire. *Families in Society-the Journal of Contemporary Human Services, 74*(10), 617-624.
Berger, R. (2010). EBP Practitioners in Search of Evidence. *Journal of Social Work, 10*(2), 175-191. doi: 10.1177/1468017310363640

Bergstrom, N. (2008). The gap between discovery and practice implementation in evidence-based practice: is practice-based evidence a solution? *International journal of evidence-based healthcare*, *6*(2), 135-136. doi: 10.1111/j.1744-1609.2008.00102.x

Brekke, J. S. (1986). Scientific Imperatives in Social-Work Research - Pluralism is not Skepticism. *Social Service Review*, *60*(4), 538-554.

Brekke, J. S., Ell, K., & Palinkas, L. A. (2007). Translational science at the National Institute of Mental Health: Can social work take its rightful place? *Research on Social Work Practice*, *17*(1), 123-133. doi: 10.1177/1049731506293693

Charmaz, K. (2006). *Constructing grounded theory: a practical guide through qualitative analysis*. Thousand Oaks: Sage.

Chorpita, B. F., Becker, K. D., & Daleiden, E. L. (2007). Understanding the common elements of evidence-based practice: Misconceptions and clinical examples. *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*(5), 647-652. doi: 10.1097/chi.0b013e318033ff71

Davidson-Arad, B., & Benbenishty, R. (2010). Contribution of child protection workers' attitudes to their risk assessments and intervention recommendations: a study in Israel. *Health & Social Care in the Community*, *18*, 1-9.

Dixon-Woods, M., Bosk, C., Aveling, E., Goeschel, C., & Pronovost, P. (2011). Explaining Michigan: Developing an Ex Post Theory of a Quality Improvement Program. *Milbank Quarterly*, *89*(2), 167-205.

Emerson, R., Fretz, R., & Shaw, L. (1995). *Writing Ethnographic Fieldnotes*. Chicago: University of Chicago Press.

Epstein, S. (2007). *Inclusion: The Politics of Difference in Medical Research*. Chicago: The University of Chicago Press.

Espeland, W., & Stevens, M. (1998). Commensuration as a Social Process. *Annual Review of Social Science*, *24*.

Espeland, W., & Vannebo, B. (2007). Accountability, Quantification, and Law. *Annual Review of Law and Social Science*, *3*.

Fischer, J. (1973). Is Casework Effective - Review. *Social work*, *18*(1), 5-20.

Flanagan, D. (2013). Evidence-Based Practice or Practice-Based Evidence? *Journal of Oral Implantology*, *39*(2), 121-122. doi: 10.1563/aaid-joi-d-11-00042

Franklin, C. (1995). Expanding the Vision of the Social Constructionist Debates - Creating Relevance for Practitioners. *Families in Society-the Journal of Contemporary Human Services*, *76*(7), 395-407.

Fraser, M., Taylor, M. J., Jackson, R., & Ojack, J. (1991). Social-Work and Science - Many Ways of Knowing. *Social Work Research & Abstracts*, *27*(4), 5-15.

Gambrill, E. (1999). Evidence-based clinical behavior analysis, evidence-based medicine and the Cochrane collaboration. *Journal of Behavior Therapy and Experimental Psychiatry*, *30*(1), 1-14. doi: 10.1016/s0005-7916(98)00035-4

Gambrill, E. (2001). Social work: An authority-based profession. *Research on Social Work Practice*, *11*(2), 166-175.

Gambrill, E. (2010). Evidence-Informed Practice: Antidote to Propaganda in the Helping Professions? *Research on Social Work Practice*, *20*(3), 302-320. doi: 10.1177/1049731509347879

Gambrill, E. D. (2002). Can't we be critical thinkers as well as hopeful? Response to Hoffman and Albers. *Journal of Social Work Education*, *38*(1), 186-196.
Gibbs, L., & Gambrill, E. (2002). Evidence-based practice: Counterarguments to objections (vol 12, pg 457, 2002). *Research on Social Work Practice*, *12*(5), 706-706.
Green, L. W. (2006). Public health asks of systems science: To advance our evidence-based practice, can you help us get more practice-based evidence? *American Journal of Public Health*, *96*(3), 406-409. doi: 10.2105/ajph.2005.066035
Ivanoff, A., Blythe, B. J., & Briar, S. (1987). The Empirical Clinical-Practice Debate. *Social Casework-Journal of Contemporary Social Work*, *68*(5), 290-298.
Lindsey, D. (1992). Reliability of the foster care placement decision: A review. *Research on Social Work Practice*, *2*(1), 65-80.
Longhofer, J., & Floersch, J. (2012). The Coming Crisis in Social Work: Some Thoughts on Social Work and Science. *Research on Social Work Practice*, *22*(5), 499-519. doi: 10.1177/1049731512445509
Mullen, E. J. (2014). Evidence-based knowledge in the context of social practice. *Scandinavian Journal of Public Health*, *42*, 59-73. doi: 10.1177/1403494813516714
Okpych, N. J., & Yu, J. L. H. (2014). A Historical Analysis of Evidence-Based Practice in Social Work: The Unfinished Journey toward an Empirically Grounded Profession. *Social Service Review*, *88*(1), 3-58. doi: 10.1086/674969
Palinkas, L. A., & Soydan, H. (2012). New Horizons of Translational Research and Research Translation in Social Work. *Research on Social Work Practice*, *22*(1), 85-92. doi: 10.1177/1049731511408738
Parsonson, B. S. (2012). The case for practice-based evidence to support evidence-based practice. *Journal of primary health care*, *4*(2), 98-99.
Peile, C. (1988). Research Paradigms in Social-Work - From Stalemate to Creative Synthesis. *Social Service Review*, *62*(1), 1-19.
Power, M. (1997). *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.
Rossi, P. H., Schuerman, J., & Budde, S. (1999). Understanding decisions about child maltreatment. *Evaluation Review*, *23*(6), 579-598.
Scott, J. (1998). *Seeing Like a State: How Certain Schemes to Imporve the Human Condition Have Failed*. New Haven: Yale University Press.
Staller, K. M. (2006). Railroads, runaways, & researchers - Returning evidence rhetoric to its practice base. *Qualitative Inquiry*, *12*(3), 503-522. doi: 10.1177/1077800406286524
Thyer, B. A. (1994). Empiricists Versus Social Constructionists - More Fuel on the Flames. *Families in Society-the Journal of Contemporary Human Services*, *75*(5), 308-312.
Timmermans, S., & Epstein, S. (2010). A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization. *Annual Review of Sociology*, *36*.
Wakefield, J. C., & Kirk, S. A. (1996). Unscientific thinking about scientific practice: Evaluation the scientist-practitioner model. *Social Work Research*, *20*(2), 83-95.

Weiss, R. (1995). *Learning From Strangers: The Art and Method of Qualitative Interview Studies*. New York: The Free Press.
Witkin, S. L. (1996). If empirical practice is the answer, then what is the question. *Social Work Research*, *20*(2), 69-75.
Witkin, S. L., & Harrison, W. D. (2001). Whose evidence and for what purpose? *Social Work*, *46*(4), 293-296.
Young, A. (2004). *The Minds of Marginalized Black Men: Making Sense of Mobility, Opportunity, and Future Life Chances*. Princeton: Princeton University Press.
Zerubavel, E. (2006). *The Elephant in the Room: Silence and Denial in Everyday Life*. Oxford: Oxford University Press.

Chapter 4
What Counts? The Quantification of Child Welfare

Introduction

In the past decade, the child welfare system has seen the proliferation of actuarial based risk assessments designed to standardize decision-making by child welfare workers. These standardized assessments have emerged in response to an array of significant issues related to clinical decision-making such as the widespread variability in outcomes among child welfare workers evaluating the same case and the profound ethnic and racial disproportionality within the system (Gambrill & Shlonsky, 2000). Scandals playing out on the front page of newspapers related to the deaths of children who were either left with their parents when they should have been removed or who died in foster care have further highlighted the need for reform in how decisions are made.

Reform efforts have focused around standardizing child welfare practices with the twin goals of (1) creating objective measures that will yield identical results for similar situations and (2) producing evidence that the child welfare system is working and cost effective. Funding priorities reflect these goals with heavy investments in the development and implementation of tools to systematize child welfare procedures and in advancing program evaluation to determine whether these efforts have been successful. As evidenced by its implementation in over 30 states, the Structured Decision Making Model (SDM) has arguably been among the

most successful among these interventions. According to the Child Information Gateway, a clearinghouse for information on child welfare practice and policy, structured decision-making is:

> An approach to child protective services that uses clearly defined and consistently applied decision-making criteria for screening for investigation, determining response priority, identifying immediate threatened harm, and estimating the risk of future abuse and neglect.(http://www.childwelfare.gov/systemwide/assessment/approaches/decision.cfm)

These assessments quantify the considerations that a Child Protective Service (CPS) worker will base her decision on, by providing a numerical score that among other things is representative of whether a child is safe, is able to remain within the home, is at risk for future harm, and the strengths of the family under investigation. The Center for Children's Research asserts this model improves the "effectiveness of the child welfare system by increasing the consistency and validity of decision-making (CRC, 1999)."

The move from child welfare decision-making policies that value the individual judgment of the child welfare worker to ones that attempt to control for professional discretion is a prominent example of the commensuration of the child welfare system. Espeland and Stevens (1998) define commensuration as the "expression or measurement of characteristics normally represented by units according to a common metric." Transforming values and ideological concepts such as safety, risk, and quality of care, into numerical scores (in the form of assessment tools) allows for the flattening of these dynamic concepts into comparable units. A long historical association between rationality and processes of classification such as ordering and numbering highlights the appeal of commensuration within child welfare (Nussbaum, 1986). Policies that promote the standardization of complex

processes demonstrate that the system is organized in a rational way and operating effectively.

Espeland and Stevens argue that the attempt to introduce statistical concepts such as reliability, replication, consistency, and validity into the course of every day life is anything but a natural phenomenon. Rather, Structured Decision-Making (SDM) and the Evidence-Based Practice (EBP) movement, of which it is a part, represent a sociological process of commensuration that requires further investigation (Espeland and Stevens, 1998). Espeland and Stevens have called for a sustained and critical examination of commensuration as it plays out across the social world because it is "symbolic, inherently interpretive, deeply political, and too important to be left implicit in sociological work" (Espeland and Stevens, 1998). To date, processes of commensuration have not been explicitly investigated as they unfold within the child welfare system.

This paper applies a sociological lens to analysis of the Structured-Decision-Making (SDM) model in order to fill a gap in our knowledge about processes of commensuration within it. I ask: What are the intended and unintended consequences of using actuarial-based risk assessments to frame decision-making in child welfare? How does state policy interact with the risk assessment to shape case trajectories? I argue that in the service of avoiding type 2 errors, type 1 errors may be more likely to occur. An investigation of the way that a 'trust in numbers' has come to organize how we approach social problems like child maltreatment (Porter, 1995)

In this paper, I briefly review the extant literature on commensuration and quantification. Following this, I examine the intended and unintended consequences

of transforming child welfare decision-making into an actuarial-based task in the state of State A. This paper concludes with a discussion of the implications of this research for child welfare decision-making and proposes policy recommendations to improve practices related to utilizing actuarial-based risk assessments themselves.

Commensuration as a Social Process[1]

Constitutive and Powerful

While a world driven by numbers (market values, market shares, statistics, prices, cost-benefit ratios, college rankings, pay) may seem natural and inevitable, sociologists who study commensuration understand this process to be a deliberate approach to the organization of social life (Espeland, 1997; Espeland & Stevens, 1998; Espeland & Vannebo, 2007a; Porter, 1995; Radin, 1996; Zelizer, 1985). In fact, the way we arrange and interpret our world has become so defined by numbers that the values and social ideas embedded within these arrangements have become all but invisible (Espeland and Stevens, 1998). Espeland and Stevens (1998) give the following example of this taken for granted process noting "commodification has become so naturalized that it is hard to construe the value of some goods in forms apart from price" (Espeland & Stevens, 1998, pg. 5). The more a commensurate process becomes integrated into every day life then the more unnoticed it is likely to become (Espeland & Stevens, 1998). As a result, investigating processes of commensuration is inherently a task of making the invisible visible, a process of uncovering how social concepts have come to be represented and contained within numerical ones.

Commensuration acts in many ways: to simplify complex concepts so that they can be understandable; to constitute relationships between abstract ideas

and/or tangible objects; to create formal procedures for accountability; to produce efficient procedures and policies, and at its most abstract, to provide the contours that define states and their activities. One example of commensuration includes the process of transforming people from diverse countries with distinct sets of cultural identities and political histories into one category (Hispanic, Arab, or European). This act renders groups who understand themselves to be radically different from one another into a single entity, identifiable as the same (Porter, 1995). Another involves the development of physician and hospital report cards to create a common metric for evaluating abstract concepts like safety, value, and skill (Espeland & Vannebo, 2007b). Report cards are intended to make it easier for patients to make decisions about which doctors to seek care from while simultaneously holding the medical community accountable for patient safety and quality of care. A third example involves efforts to legitimatize the work of housewives by assigning value to their domestic efforts (Espeland & Stevens, 1998). Assigning economic values to housework is shorthand for communicating that this work is actual labor with a monetary benefit to the family. Through commensuration concepts that are thought to be intangible are given a concrete expression through numbers. In the process, they may be granted new meaning (Espeland, 1997; Espeland & Stevens, 1998; Latour & Woolgar, 1986; Porter, 1995).

Commensurate processes have many benefits. Giving complex concepts a shorthand is a way of simplifying difficult ideas, creating efficient paths for making decisions or responding to information that would otherwise be overwhelming. (Espeland and Stevens, 1998; Espeland and Vannebo; Stockey & Zeckhauser, 1978). Exhaustive research in cognitive science and psychology demonstrates how schemas

and strategies are necessary for managing and organizing multiple data points, without which it would be all but impossible to go through every day life (Espeland & Stevens, 1998; Tversky & Kahneman, 1974, 2002). By recasting complicated information into easily understandable units or numbers, processes of commensuration offer an important strategy for negotiating the world and focus our attention on the information that is important to consider while leaving out information that might be unnecessary. The technological advantages of the digital age align naturally with what is required to commensurate complex ideas with computers providing the means for analyzing any number of issues through complicated formulas and algorithms.

Commensuration also has the potential to democratize (Espeland & Stevens, 1998). By counting, measuring, or depersonalizing the way information is handled and decisions are made, commensurate processes offer systematized and consistent ways for making sense of complicated or opaque processes. Using numbers whether in the form of statistics, prices, or other measures of comparable worth offers a way to make information accessible to the public and creates standardized ways for analyzing information and making judgments (Espeland and Stevens,1998). In this way, commensuration can install transparency and curtail decisions that are made based on subjective assessments, relationships, favors, or back room deals. By allowing the public to review and evaluate the decisions of those in power, common metrics create a mechanism by which the actions of those in a position of authority can be limited (Espeland, 1997; Espeland & Stevens, 1998; Espeland & Vannebo, 2007b)

At the same time, standardized measures function to limit discretion and local knowledge by dictating the processes by which decisions are made, also making them appealing instruments for those in power to implement (Espeland, 1997; Espeland & Stevens, 1998) Because commensuration can be used to both shore up and constrain authority, opposition to its practices can be found among both liberals and conservatives alike (Espeland, 1997; Porter, 1995).

Commensuration can fundamentally alter the meaning of the things that they are trying to measure, standardize, evaluate, or compare (Espeland and Stevens, 1998). Art given a high monetary value may draw people to a particular painting that they do not like but chose to buy anyway because of its worth, changing the relationship between personal taste and ownership or personal taste and value. The power of counting and measuring" can even create new social categories and relationships where they did not previously exist (Espeland, 1997; Espeland & Stevens, 1998; Latour & Woolgar, 1986). Porter asserts that statistical concepts have not only defined the idea of society but also authored it. (Porter, 1995) As phenomena that were greater than individual acts such as suicides, birth rates, accidents, crimes, were able to be counted, patterned, and analyzed, new concepts were needed to understand them, giving rise to the identification of society itself (Porter, 1995, 1986; Espeland and Stevens, 1998; Durkheim, 1951; Hacking 1990).

As Porter, Espeland and Stevens all note "numbers become self-vindicating (Porter, 1995, p. 45). Their importance is reified as commensurate processes become institutionalized (Espeland and Stevens, 1998; Porter, 1995). Law school rankings have transformed not only which students are admitted to institutions but how the institutions themselves behave to compose their class giving out more merit based

scholarships to those with high LSAT scores or weighting LSAT scores higher than other criteria (Espeland & Vannebo, 2007b). Once a measure or metric is established as important, it then comes to define how people act in response. In turn, people's behavior grants continued legitimacy to the commensurate process. Commensuration then is not just imposed on people, to hold power it must, at least partially, be accepted as a legitimate way to organize life (Espeland & Sauder, 2007).

Accountability and Commensuration

One way that commensuration produces new relationships is by linking quantified processes to the measurement and evaluation of performance (Power, 1997; Porter, 1995; Ranson, 2003; Espeland and Vannebo, 2007). Indeed, the connection between accountability and commensuration cannot be overstated. In the last 40 years, efforts to define outcomes and rate the quality and efficiency of performance have transformed the methods with which work is conducted (Espeland & Vannebo, 2007b; Porter, 1995; Power, 1997). Reliance on quantitative assessments to do and assess work, symbolizes the emergence of what Power has termed "the audit society' (Power, 1997). Power understands the audit society to replace trust in professionals with trust in the procedures and organizations that monitor them (Power, 1997).

Quantitative assessments have become predominant means by which professionals and institutions are held accountable. Porter conceptualizes this shift to one where work is guided by the predominance of "mechanical objectivity" (Porter, 1995). Mechanical objectivity "is knowledge whose authority is based on compliance with explicit rules" (Espeland, 1997). In its formality, mechanical objectivity replaces personal knowledge and expertise with a "trust in numbers"

(Porter, 1997). Mechanical objectivity fills a need for creating procedures for decision-making that are depersonalized, able to be coordinated across multiple actors and domains, to be transparent, impersonal, clear, and fair. Numbers are seen as a neutral and therefore offer an ideal way to meet these criteria (Espeland, 1997; Espeland & Stevens, 1998). However, multiple theorists have documented how numbers are anything but impartial and objective in the way that they are employed and given meaning (Bowker & Starr, 1999; Epstein, 2007; Espeland, 1997; Espeland & Sauder, 2007; Espeland & Stevens, 1998; Espeland & Vannebo, 2007b; Hacking, 1990; Lampland & Starr, 1999; Latour & Woolgar, 1986; Porter, 1995; Timmermans & Epstein, 2010). Not only is effort required to determine how numbers and procedures will come to symbolize concepts but the processes by which they are given meaning are often obscured (Bowker & Starr, 1999; Espeland & Stevens, 1998; Lampland & Starr, 1999; Latour & Woolgar, 1986).

Porter argues that mechanical objectivity emerges under specific social conditions, particularly where trust is at issue (Porter, 1995). In the United Kingdom, nineteenth century actuaries resisted standardizing their methods, claiming that discretion was key to performing their role. n contrast, after the Great Depression, American accountants adopted standardized practices when the political climate made anything but uniform methods untenable as a way of doing business (Porter, 1995). Distinct attempts to standardize the same phenomena and practices have unfolded differently in distinct contexts highlighting this process to be social rather than technical.

Mechanical objectivity is most often valued when the public does not trust official decisions and those who make them, when there is conflict about what to do,

when decisions are open to review, and when a common language is needed to traverse social and physical distance (Porter, 1995, Espeland and Vannebo, 2007b). Decisions guided by quantitative processes, Porter asserts, are then most likely to be practiced by weak elites in political democracies (Porter, 1995). Mechanical objectivity satisfies the needs of bureaucrats who must defend their decision-making to others and who must demonstrate that their actions are fair and impersonal (Porter, 1995). As Espeland summarizes:

> This is why the United States, with its long tradition of distrusting elites and experts, relies so heavily on rules for constraining official discretion. It also explains why the United States has led the way in producing and promoting such technologies as cost-benefit and risk analyses, educational and intelligence testing, opinion polling, survey research, and quantitative social science (Espeland, 1997).

This historical and political moment, where quantified processes are the preferred method for decision-making and accountability is a particular one, guided by certain political realities and cultural norms. It is not the only or inevitable way to organize our approach but right now it is the dominant one.

A Moral Process

Not only is commensuration a political process, it is also a moral one (Strathern, 1996; Espeland and Vannebo, 2007b). When numbers are used to determine the answers to complex social problems, they encode within them ethical judgments and values. Espeland and Vannebo (2007) detail how sentencing guidelines "mediate the distribution and meaning of punishment such that justice is rendered to abstract categories of persons rather than particular individuals" (Espeland and Vannebo, 2007). How justice is meted out, what crimes are considered worthy of more extensive punishment and then what happens as a result are issues

embedded within quantitative responses to social problems such as sentencing guidelines. The difference in mandatory minimums for possession of crack/cocaine versus cocaine and the attendant consequences for who is incarcerated and for how long, exemplify the ways that social ideas about class and race can become obscured by numbers meant to be neutral. Even the desire itself to make processes efficient and decision-makers accountable reveals specific cultural mores (Espeland and Vannebo, 2007; Porter,1995).

The Structured Decision-Making Model in State A

Because commensurate processes contain within them moral and ethical values, they are particularly important to carefully unpack. Deconstructing what is and is not emphasized within quantified or mechanized assessments offers clues as to how social problems are bounded, understood, and then ultimately dealt with. Such an understanding is an important and often overlooked part of social work research and child welfare practice. Without understanding the embedded values within structured decision-making practices (and other decision-making protocols), we cannot fully evaluate whether such a response makes sense. A sociological investigation of the Structure Decision Making model has the potential to reveal how SDM itself structures epistemological ideas about what constitutes maltreatment and the criteria for subsequent decisions throughout a child welfare case.

This article examines commensuration in the child welfare system by investigating the Structured Decision-Making Model as it is practiced in one state. A close reading of how one component of the SDM, the risk assessment, is constructed and then utilized offers insight into the consequences, both intended and

unintended, of quantifying the decision-making process within Child Protective Services (CPS).

While there are multiple components to the SDM, this article only examines the construction and use of the risk assessment[2], which is arguably the most critical part of the SDM as it is practiced in a state that makes it central to child welfare cases. Tied to a number of decisions such as: the level at which a case is categorized; [3] whether a parent or caregiver is placed on the Central Registry for Abuse and Neglect (a centralized list of offenders of child maltreatment searchable by employers and others); the intensity of services a family will receive and whether those services will be voluntary or mandated; and the frequency of contacts with the Child Protective Services; the risk assessment defines the shape a case will take and therefore exemplifies how a process of commensuration impacts child welfare decision-making. Use of the risk assessment is mandated by the state legislature further underscoring its importance

The risk assessment

In order to examine how the risk assessment is used in State A, an understanding of its components is required. Based on principles of actuarial risk assessment, this tool attempts to determine the likelihood that parent(s) or caregiver(s) will neglect or abuse their children in the future. The content of the risk assessment varies state by state but the format remains the same. The risk assessment discussed in this paper contains two sets of inventories one for neglect and one for abuse. Factors associated with abuse and neglect such as (but not limited to) parental history of abuse, history or presence of domestic violence, number of children in the house, age of caregiver, and previous involvement with

child protective services comprise the scales. Points are assigned to each item that the caseworker endorses. For example, if a worker agreed with the statement that the "parent has a mental health problem," one point would be assigned. Point values for each statement are weighted to reflect distinct levels of severity and impact. The statement: 'Primary caretaker involved in harmful relationships' contains one point for endorsing 'harmful relationships and one domestic violent incident' and two points for endorsing 'multiple domestic violence incidents.' At the completion of the assessment, the total points from the neglect inventory and the total points from the abuse inventory are calculated separately. The highest number from either inventory reflects the score used. A parent could score relatively low on the abuse inventory (3) but high on the neglect inventory (6) and the score for the parent would still be (6). The numerical value of the risk assessment represents whether a family is at an Intensive, High, Moderate, or Low risk for recurrence of abuse or neglect with higher numerical scores correlating with higher risk categories. A finding of Intensive on the risk assessment indicates the need for an immediate petition to the court for removal.

Risk assessments are developed using validation studies to identify characteristics of families that indicate the potential for future maltreatment. Validation studies determine the most influential risk factors for maltreatment, the combination of factors that are most likely to contribute to abuse or neglect, and the appropriate statistical weights and measures necessary to create a risk profile for families using this information. Measures for the validation studies are collected using CPS data about the characteristics of families who have been substantiated for

abuse and neglect as well as information about which families have been re-referred for child maltreatment after an initial investigation and/or substantiation.

(http://nccdglobal.org/sites/default/files/publication_pdf/nc_risk_asst_validation_final2009_2.pdf).

The Children's Research Center expresses enormous confidence in actuarial risk assessment noting "because these tools are products of research on the actual experience of families previously reported to the agency, it is possible to assess risk with a reasonably high degree of accuracy"(CRC, 2008). Distinguishing between the benefits of classification and predictive power, the Children's Research Center carefully emphasizes that the risk assessment is not necessarily prognostic about whether families will maltreat their children in the future. However, case actions are largely influenced by the potential for future harm as indicated by the risk level (CRC, 2008).

Beyond providing information about the risk for on-going child maltreatment, the risk assessment also guides decision-making about whether a family would benefit from services if a case is substantiated and opened. Risk levels dictate the level of intervention and oversight necessary for families. Families with a low risk level have infrequent face to face contact with their caseworker (once a month) and caseworkers make similarly infrequent connections with providers (one collateral contact per month), while caseworkers will make four visits per month and four collateral contacts for families that are deemed to be high risk.

Mandating the frequency of interactions that a family has with CPS based on their risk level is intended to distribute the time and attention of caseworkers according to the greatest need. Limited resources are then not spent on families who

are least likely to benefit from system engagement, leaving more opportunities to supervise families where the potential for continued or future maltreatment remains high (CRC, 2008).

In State A, caseworkers and supervisors have the option to override a risk level that they believe does not represent the severity of a particular situation. There are two types of overrides: discretionary and mandatory. Discretionary overrides can be made by a worker to increase the risk level to a higher category with written justification and supervisor approval. These overrides can only be used to raise the risk level. Mandatory overrides are policies that require the risk level to be raised to Intensive in the presence of certain conditions such as: a child with severe injuries resulting from the action or inaction of a parent, the death of a sibling from abuse or neglect by the shared caregiver, alleged sexual abuse when the accuser has access to the child, a child under the age of two with serious non-accidental injuries, a parent who has had another child after having their parental rights terminated. Once a case has been substantiated, the investigator has 30 days to complete the risk assessment. If a case is denied, the risk assessment does not have to be filled out and is not used to inform decision-making.

Research Methods

Sample

The data for the present study originates from in-depth qualitative interviews with 35 Child Protective Service (CPS) workers and close reading and analysis of primary source material about the Structured Decision-Making Model in State A. I recruited participants for the study using two related strategies. First, I presented the study and passed around a sign-up sheet for those were interested in

joining at continuing education trainings at a School of Social Work. I then followed up with these self-identified volunteers and scheduled interviews. Additionally, a director of a child welfare agency in introduced my research to his staff and passed around a sign-up sheet. Following an identical procedure, I contacted volunteers and scheduled interviews with those who demonstrated interest. At this area office, I also recruited participants using a snowball sample as other CPS workers heard about the study from colleagues and asked to join. Participants were offered a $25 Visa gift card to compensate them for their time.

Interviewees ranged in age from 24-62. Participants were comprised of 17 White women, 8 African-American women, and 6 White men. Workers interviewed had a variety of experience. 42 %(15 out of 35) had worked in their current position less than 2 years, 22% worked in their current position between 2-5 years (8 out of 35), and 20% (7 out of 35) had over 5 years of experience on the job. All participants were front line Child Protective Service workers tasked with either making decisions either about substantiation and severity of child maltreatment (Investigators) or, in the case of substantiated cases whether a family was complying with treatment and recommended services (Treatment Workers). For decisions about the latter group, Child Protective Service workers were responsible for either discharging or escalating cases as families wound their way through the child welfare system

Interview Content

The interview protocol for this study was designed to ground data about participants beliefs through their behaviors (Charmaz, 2006). Semi- structured and open-ended, the protocol focused on eliciting information about their every day

work practice. I collected general information on what a typical day looked like, how a participant chose their profession, and the challenges of the job as well its pleasures. Following this initial discussion, I asked participants "to walk me through" a series of cases from the time they opened them to the time they closed or transferred them (Weiss, 1995). These cases included their most recent neglect case, a challenging case, and a case that were they felt good about the decision they had come to. By collecting data on cases across a continuum of affective states and challenges, I was able to get a rich sense of how a worker conceptualized their role, what strategies they regularly engaged in as child welfare workers, and what they understood to be 'typical', 'hard', and 'easy'.

After collecting these accounts, I turned the focus of the interview to the Structured Decision-Making (SDM) tools (if they had not yet mentioned them at other points in the interview). Once again, I collected information on attitudes through questions about behaviors. I asked workers how they used these tools in their every day practice and asked each worker to 'walk me through' through the Safety, Risk, or Reunification Assessments for any or all of the cases they described. These questions allowed me to collect data on the weight that these assessments were given in the final disposition of the case and how they influenced decision-making about legal or social service involvement.

Throughout the interview, I followed up every question with probes specific to the conversation. Common probes focused around further elucidating the logic at decision-making points and case detail. Interviews ranged from 60-180 minutes with a median interview time of 90-100 minutes. Interviews were audio taped and transcribed. I conducted interviews at neutral locations such as local libraries and a

community center with offices available for sign-out. Interviews also took place in an available room at one area Child Protective Service office. All names of offices, agencies, and people are pseudonyms. (Bosk, Same Difference: Standardized Decision-Making in Practice, In Progress)

Data Analysis

Following grounded theory protocols, I began analyzing data at the same time that I was collecting data(Charmaz, 2006). Using an iterative approach, I began coding interviews without any a priori ideas or assumptions about what themes were contained within them. This open coding strategy was first applied to individual interviews using Nvivo 10 qualitative data analysis software. (Emerson, Fretz, & Shaw, 1995). Codes were selected based on the emergence of patterned regularities in the transcripts. Once I identified core themes, I then coded interviews at the aggregate level. After a codebook had been developed, two Research Assistants analyzed the data deductively based on the themes and categories I initially identified to ensure inter-rater reliability. The data presented below represents larger themes found in the data, except where noted. Quotes from the data are edited for clarity (e.g. repeated "Ums" removed), but no words have been changed or re-ordered.

Why Here

State A offers an ideal site in which to conduct this research. Currently, State A's legislature mandates the use of the Structured Decision-Making Model (SDM) for child welfare decision-making in all casework. In addition to the SDM, State A has also adopted a category system where Child Protective Service workers are

required to rank a case based on level of severity from 1-5. The ranking system further standardizes caseworker action and discretion by specifying the specific protocols that follow each ranking. Case categories are explicitly linked to the risk assessment (RA) with scores on the RA correlating with the case category. The chart below outlines the category system and their connecting designations from the risk assessment: (Bosk, Same Difference: Standardized Decision-Making in Practice, In Progress)

State A Child Protective Services Category System

Category V	Following a field investigation, the department determines that there is no evidence of child maltreatment or is unable to locate the family. **No action taken**.
Category IV	Cases where there is no preponderance of evidence that neglect or abuse has occurred. If the risk assessment demonstrates any risk **Voluntary Community Services are Recommended.**
Category III	A preponderance of evidence that child maltreatment has occurred. The risk assessment scores as low or moderate. If the family does not voluntarily participate in services or does not progress CPS may reclassify as a Category II. **Community Services are Needed. Moderate Risk**
Category II	A preponderance of evidence that abuse or neglect has occurred. There is a high level of risk for future harm as indicated by the risk assessment. The perpetrator is listed on the Central Registry for Abuse and Neglect. **Child Protective Services are Required in Conjunction with Community Services. High Risk**
Category I	A preponderance of evidence that child maltreatment has occurred with the risk assessment scoring as high or intensive. Child Protective Services determine that a court petition is either needed or required. Perpetrator is listed on the Central Registry for Abuse and Neglect. **CPS petitions for the immediate removal of child(ren). Services Provided by Child Protective Services or Foster Care in Conjunction with Community Services. Intensive Risk**

Avoiding Type 2 Error leads to Type 1 Error

In State A, the policies related to the Structured Decision-Making model as well as the items on the risk assessment itself appear designed to avoid a Child Protective Service worker making a type 2 error: leaving a child in a home when s/he should have been removed. Discretionary overrides that permit a case worker to raise the risk level as well as mandatory policy overrides for specific situations function to ensure that there are multiple mechanisms that enable a removal to take place or dictate that it does when certain conditions are met.

The fact that overrides can be made to raise but not lower the risk level signal the explicit concern for avoiding one set of errors in judgment: those that involve having a child remain in a situation that is unsafe. The consequences for making this kind of error cannot be overstated. Children have died when Child Protective Service workers have assessed a home to be safe when it is not. The moral and organizational incentives to avoid a child suffering or dying as the result of a flawed assessment are clearly enormous. High profile fatalities often lead to very public excoriations of the Department of Health and Human Services in the forms of blistering media coverage, public outcries for change, and state hearings to uncover what went wrong. Indeed child welfare policy itself is enormously reactive to scandals that most often involve the death of children who have come into contact with the child welfare system(Gainsborough, 2010). Gainsborough has argued that responses to child fatalities are actually the single driving force for making changes to existing policies or creating new ones within the child welfare system (Gainsborough, 2010). It makes sense, then, that actuarial-based risk assessments as well as the policies surrounding their use would be designed to avoid making this

kind of error. However, it may be that in the desire to avoid committing a type 2 error that the conditions arise under which more type 1 errors (removing a child from a home when they should have been able to remain in place) are made. I argue that a combination of prescribed actions from the Department of Health and Human Services and the construction of State A's risk assessment make type 1 errors more common.

Removals when they are unwarranted can be absolutely devastating to children and their caregivers. For young children, disruptions in care may profoundly alter relationships with life long consequences for their emotional, social, and cognitive development (Bowlby, 1973, 1979, 1980). For caregivers, extensive involvement with the child welfare system comes with costs related to time, money, energy, self-esteem, and wellbeing. Competent legal representation is often necessary to facilitate a reunification and access to these services may not be available to all who require it. Caregivers and their children, often obscured from the view of those who made the original decision, hold the burdens associated with an unwarranted removal. As a result there is not the same organizational incentives to limit these mistakes as they are not as visible to those who made them, the media or the public. The consequences of type 1 errors make it equally important to understand how and when they are made within the decision-making process.

For some parents and caregivers, risk assessments that are rated higher than they should be do not result in the removal of their children but do result in their placement on the Central Registry for Abuse and Neglect. The Central Registry for Abuse and Neglect is a state list of people who have had serious charges of child maltreatment substantiated against them. This database is publicly searchable and

those who are placed on it are prohibited from holding a job in which they come into contact with children, from participating in activities in which multiple children are present (such as attending school functions or accompanying their children on field trips), or from having a child placed with them by Child Protective Services in the future. Many people who land on the registry are at risk for loosing their job so there are serious consequences that can last long after a maltreatment case has been concluded. This next section considers the ways in which the SDM risk assessment in State A may mean that in the service of avoiding a type 2 error, a type 1 error is more likely to be committed instead.

Demographic Factors

The risk assessment contains a mix of demographic and dynamic factors related to risk for child maltreatment. Demographic factors include the number of children in the home as well as the age of the children. Focus on demographic factors means that it is possible for a risk assessment to come out as high through the presence of static indicators rather than the behaviors of the caregiver in question. Angela describes how this can occur:

> A: Well sometimes with the outcome of the risk assessment I don't agree with; like 'cause you know when it comes to high and it's open, it's gonna be a Category 2. And sometimes I feel like it shouldn't be 'cause sometimes it's just for having four or more kids in the home…
>
> E: So like that might be something that throws it over the edge?
>
> A: That throws it over the edge and it shouldn't; like or either there'll be like there were four or more kids in the home and they had…and the children had mental health issues. Then that throws it over the edge to make them be a Category 2 and put them on the unfit list; even when it's like a first-time offense.
>
> E: What do you do in those situations? 'Cause it sounds like…
>
> A: There's nothing you can do

Cases that are categorized as 2 or higher mean an automatic placement on central registry for the parent or caregiver under investigation. As it is currently constructed, the State A risk assessment weights having more than 2 children, children under the age of 7 and a previous child maltreatment complaint (not necessarily a substantiation) as factors that will raise the risk level for a family. Questions about the number of children in the home are asked twice (once on the neglect inventory and once on the abuse inventory) with the risk level being raised for more than three children on the neglect side. Workers report that families have often been penalized for having young children or more children rather than on relational factors or behavioral factors related to the complaint.

Demographic items that seem fairly straightforward on their face and without any controversy as to their relationship to child maltreatment such as N2. *Number of prior assigned neglect complaints and/or finding*s and A2. *Number of prior assigned abuse complaints and/or findings* can be more problematic in practice. The assumption behind including them in the risk assessment is that prior history is a predictor of future behavior. The more complaints a parent or caregiver have had against them in the past, the more risky the current situation is likely to be and the greater the probability of future maltreatment. By collecting measures on the number of complaints rather than on the number of substantiated complaints, the risk assessment treats the suspicion of maltreatment by others and the presence of confirmed incidents of maltreatment as the same event or at least as events that predict the same amount of risk.

Investigators report that the reasons for multiple complaints vary and are not necessarily indicative of the parent or caregiver's past actions. Parents engaging

in custody battles can use the Child Protective Service System as a way to bolster their case for receiving custody. Filing an anonymous complaint with CPS is also a way of settling scores between neighbors, acquaintances and of exacting revenge when relationships between parents change (separation, breakups, divorce) and when new people (girlfriends, boyfriends, biological parents of half siblings) enter the picture. During my fieldwork, investigators estimated that somewhere between 50- 80% of their caseload was dedicated to evaluating reports that were clearly false and personal in nature. Of these false complaints, investigators estimated that almost half of these risk assessments scored as high.

Because unsubstantiated complaints are treated the same as substantiated ones in child maltreatment investigations, they can and do inflate the risk score. As Ashley describes:

> …You'll have some families that the neighbors will continually call on them or there'll be like one RP [Reporting Person] that continually calls on them and makes up a new allegation every time and we have to go out and investigate and deny the case repeatedly. But let's say they have four kids and one's young. So then they got a bunch of previous CPS history, the young kid and the four kids in the home, they're already at a very high-risk level. So even if it's a small issue because of all those risk points, there's a very good chance they'll be put on Central Registry and Cat 2. I think that's one issue I do have with the abuse assessment is those three categories. See the previous CPS history, it should be not just assigned cases but substantiated cases.

Similarly, Diane observes that just the presence of previous complaints is enough to move the risk assessment from low to moderate:

> I mean I think sometimes when there's prior history, you know you have to put in prior history, and sometimes that can elevate it more than it would need to be, because maybe the situation that you have right now isn't severe, but because of the history it's, you know it will elevate the risk assessment. So those times, you know you would maybe have a disparity between the two; like this really doesn't warrant a high risk level, but it's coming out high because they've had two prior cases.

Literature has demonstrated that CPS allegations and investigations occur not just at the individual level but at the neighborhood level as well (Claudia J. Coulton, Korbin, & Su, 1999; C. J. Coulton, Korbin, Su, & Chow, 1995; Garbarino & Crouter, 1978; Garbarino & Kostelny, 1992; Garbarino & Sherman, 1980; Zuravin, 1989). Communities where most residents are of low socio-economic status as well as communities of color are more likely to come under the official scrutiny of Child Protective Services than those of high socio-economic status and those that are mostly white. Research on disproportionality in the child welfare system has demonstrated that maltreatment is not more likely to occur in communities that receive more scrutiny but that the surveillance itself predicts involvement with the child welfare system (Fluke, Yuan, Hedderson, & Curtis, 2003). These issues come to a head in questions designed to capture suspicion of families rather than actual incidents of maltreatment. In cases where maltreatment is substantiated, previous scrutiny by the child welfare system, even when it was unfounded, can be counted against a parent. The risk assessment encodes the idea that contacts with the child welfare system in any form raise the likelihood of future maltreatment.

The Presence of the Past: How Personal History Impacts Current Assessments

Playing out within the risk assessment form are larger discussions about the role that a parent's experience of maltreatment has to the current complaint. Question A5 asks workers to record whether either caretaker was abused or neglected as a child. If the question is endorsed 2 points are added to the risk assessment. Workers assert that in combination with other demographic factors such as age of the child and number of children in the house that adding past history into the mix can be enough to move caregivers into a High risk category and on to

the Central Registry for Abuse and Neglect without regard to other factors that should influence the case such as such as the severity of the substantiated maltreatment and how those under investigation understand and are addressing the problem. Jennifer asserts:

> I mean there's cases all the time that score out as high, you have parents participating throughout the whole thing… You know one of the biggest questions we have to ask in policy is; does the family... have you ever been abused or neglected as a child? They self report that. Well again, if you clicked yes, parent self reports they were sexually abused as a minor, you know again that just raised their risk level, but that doesn't mean that they're at [risk]... it could've been over something of a dirty house you know.

Jennifer notes that history of parental maltreatment might not impact the current complaint in a meaningful way. However, as the item is currently written, there is no way to have a more sophisticated understanding of how a parental history of maltreatment as a child affects the current situation.

It is reasonable to assume that rather than affecting all families in the same way (raising the likelihood that a caregiver will maltreat a child) that there are multiple ways that a history of maltreatment could impact caregiving. A parent could construct her approach to child rearing in opposition to what she herself experienced or a parent could repeat what she herself had learned. The experience of child maltreatment itself is suggestive and likely impacts how a parent parents but it is not determinative in any particular direction. Further, this question, as it is phrased, does not distinguish between caregivers who have received treatment for the negative effects of maltreatment and those who have not. In doing so, the inventory understands a history of maltreatment uniformly to be both negative and static, collapsing events that are dynamic and multilayered into a singular meaning.

Questions about the past whether related to a parent's experience of maltreatment, previous complaints, or past incidents of maltreatment are phrased only to capture the existence of these events rather than determine their meaning in the present. In this way, events that are dynamic and multi-layered are smoothed into a singular and flat occurrence.

Lauren recounts how a risk assessment scored as high even when her investigation revealed no risk to the children:

> There was a family who had previous CPS history a lot of it, which bumps it up on the risk scale. You know here was previous domestic violence, previous substance abuse, and the father I think was abused or neglected as a child. So the risk level kind of scored through the roof, but the allegations we got were that the children they didn't have food or clothing. And I went to the house and everything was perfect. All the kids, they didn't disclose anything concerning and there was plenty of food; they had tons of clothes. So it was kind of bogus allegations I think, so obviously I wasn't going to substantiate the family, but the risk level did score through the roof because of everything that had happened in the past...(short laugh)…so and you can't bump it down a risk level, you can only bump it up so in those cases you just got to leave it as a high risk and deny it.

While a negative outcome for the family is prevented from occurring because the case was denied, the family would have had a very different outcome if Lauren had found cause for maltreatment that needed to be addressed but was not severe. This leaves little room for workers to address a maltreatment complaint that they might need to substantiate but where the severity of maltreatment does not match the risk assessment.

While there is a clear link between a parents' history of maltreatment and maltreatment of their children, it is by no means a fixed one Not all people who were maltreated as children go on to maltreat their own children. In fact, studies had shown that transmission of abusive behaviors varies between 18-40% for all abuse types (Gara et.Al. 2000; Pears & Capaldi, 2001; Elwell, & Ephross, 1987; Kelly,

1990, Newberger, Gremy, Waternaux, & Newberger, 1993; Oates, et al., 1998). Emphasis on the past without the ability to consider its meaning in the present takes a determinative view of the relationship between events, embedding theories about the cycle of violence within the risk assessment itself.

When analyzed collectively, one of the main critiques that Child Protective Service workers in State A have about the risk assessment is the way that it can penalize caregivers for being 'a victim of circumstance.' Because caregivers can end up as a high-risk case simply from a combination of demographic and historical factors, investigators report that, at times, cases escalate when they could be handled in other ways. Demographic items that document facts or events are seen differently by workers than those that encode agency or behaviors. The former are viewed to be more 'unfair' by workers because they may or may not be related to whether the substantiated charge merits the actions that occur as a result. Child Protective Service workers assert that having little room for both interpretation on the risk assessment and then the resulting case actions means that sometimes cases unfold in ways that they feel are unjustified.

Casey explains how the presence of a certain amount of risk factors unmediated by more nuanced reflections on the impact these risk factors are having on a case can lead to situations were workers are required to file petitions to remove children when they do not believe this action is warranted:

> It was a neglect case and that's one of the ones I ended up having to do in-home ward with the courts. Mom is, I believe between 24 and 26 and she has mental health issues of her own, and she comes from a very difficult relationship with her parents; a difficult relationship with her family. I think her mother is also undiagnosed bi-polar so I can only imagine how bad they, you know don't get along.And she has 3 children; an 8-year-old, a 5-year-old and a 1-year-old, I think she's like 13 months. And she was living with the father of the two youngest and for some reason she decided

> to let the oldest boy live with her parents and so she was only living with the two youngest and the father.
>
> And it was a domestic violence relationship and the house was in disarray partially because the landlord hadn't cared for, like the water main broke. So between that and a little bit of them knocking around and throwing things and her being, her mental health issues. When the police came out for a domestic violence incident they filed because the house was dirty.
>
> I came out and she ended up getting evicted from the house within a couple of weeks because they weren't paying rent, I guess because the landlord wasn't fixing the issues. And when, you know I looked at the case because of mom's mental health issues, because she has three children and one of, her youngest baby was special needs. She was in a domestic violence victim, she had CPS history as a victim; all of these things made her have a high risk and made us have to go to the point where we had to file in-home.
>
> But she was completely willing to work with services. She was so happy to have someone to help her get out of the domestic violence situation. To help her get her own housing because she's always relied on her mom or a man and this was her first opportunity to support herself and we were helping her and she went to counseling and you know, she went with Families First and this was all before we filed. And just because her risk assessment was so high; we had to file.

In this case, Casey believes that while multiple risk factors (the mother's mental health, the domestic violence, housing situation) are risk factors that need to be addressed, the outcome for the case is unnecessarily punitive. The risk assessment does not take into account the relationship between the risk factors at hand and the behavior of the parent under investigation. Casey observes that the mother in this case is working hard to attend to the range of issues that have brought her to the attention of the child welfare system. In her estimation, these efforts are enough to mitigate the higher risk assessed by the risk assessment. Casey sums up:

> And I totally agreed with the allegations and that somebody needed to you know address her mental health, address why she is doing this, and give her parenting classes to address the fact that she can't parent her kid and she just wants to medicate them to deal with them.
>
> But I didn't think that it was to the level that we had to take her job away and make it so she couldn't care for her four children, because then if she can't, you know that's just going to make the whole situation worse. And, I mean, I don't know I just felt that the way that, if she is a risk to children I

> feel like she should be a risk to children because of the actions and decisions she made, not because of the things that happened to her

By tying the trajectory of the case solely to the presence of risk factors themselves, without taking into account how the parent is responding to them may create the conditions under which more severe outcomes are likely to occur. Even worse, by jeopardizing the employment situation of this mother, placement on the Central Registry might actually make it more likely that child maltreatment will occur in the future by creating the conditions for physical neglect.

Keeping it Simple Is Not Less Complicated

Kelly finds that the lack of nuance that Casey refers to on the risk assessment makes it difficult to complete it accurately:

> It's either – it is this or it isn't this. There's no grey area you know, so it's like, and that's what makes it hard going through this. You know it's like; am I gonna want to put another thing on the risk assessment or is it not to that point where I really need to put it on the risk assessment? Its just not so straightforward, you know…

Like all standardized assessments, the risk assessment in State A is actually designed to mediate the gray areas of cases. By using the presence or absence of identified risk factors to sort families into assigned categories, cases that are not straightforward are intended to be made more so. One of the goals of the risk assessment is to create a more objective reading of the facts of the case. Quantifying the number of risk factors and then assigning each of these a weighted score explicitly removes the context that accompanies every situation. The intent is to make decisions fairer by grounding judgments in the latest research and eliminating the subjective elements of decision-making. Ironically, taking away the potential for contextual assessment also contains the potential to make decisions more unfair. Without the opportunity to account for nuance or complexity, cases that are

fundamentally unalike are treated as the same. William describes how flattening the complexity of cases can actually invite more complication:

> We have kids that beat up their parents. I know we're here to protect the kids, but we, when I in foster care we had kids that were [in] 7 and 8 old in residential placements. They come back to the community they're out of control. They're parents are sitting on them holding their hands down because they've busted everything up and they're hurting everyone in the house, and they don't know what else to do, and the child may have a bruise on his wrist from being held down. That's a bruise, that's substantiation.
>
> That again, we're humans there is not anything that fits perfectly. Those are the things we can't change. I can't [change] the fact this kid has a bruise. Just because of that, it's going to be open, period. We need the discretion to say; how did that bruise get there? Was mom protecting the child? Was mom protecting another child from a child? Those things go on.

William asserts cases that may seem straightforward on their face (e.g. corporal punishment by a parent that results in a mark, which is not only grounds for a mandatory substantiation but also against the law) in practice are anything but clear. Processes of commensuration often treat discretion as suspect. An unintended consequence of abstracting concepts into numerical forms can be an outcome that violates the principle of the rule.

While the risk assessment is designed to protect families from the harmful consequences of subjective assessments that are incorrect, rooted in bias, or based on erroneous beliefs; the flip side of removing discretion from casework is that when the assigned case trajectory does not match the case dynamics, caseworkers have almost no recourse to redress the situation. Kelly sums it up:

> And that's where the whole grey area, you know it's for a question like that, it depends on the family. Because like I said you could walk into a home…and it's exactly as it appears [on the allegations]. They are absolutely not willing to take care of anything. But here's a mom that loves her child more than anything and here's a daughter that loves her mom you know, but one bad thing happened and it snowballed into a mess. And it's just an unfortunate situation for everyone.

Kelly is discussing a case where the actuarial risk assessment suggests one course of action but where its context invites a very different conclusion than the numbers draw. As Kelly notes, the meaning of any item on the risk assessment really depends on the family.

Lack of nuance begins on the risk assessment itself, in the way that questions are phrased, and continues through how State A) ties policy to the outcome of the risk assessment. By correlating the result of the risk assessment with the case category, State A does not see a role for assessment outside of the form once a case has been substantiated. When it comes to cases that require an immediate petition for removal, failures of the risk assessment are assumed to occasionally occur. The mandatory overrides are meant to account for situations where the risk assessment does not capture the full meaning of a case. Yet, no such accounting is made for scenarios where the risk assessment understands the situation to be more severe than a caseworker believes it to be. In this way, state policy views the risk assessment to be infallible when it comes to overestimating risk and fallible when it comes to underestimating it.

Two Sides to Every Story: Intimate Partner Violence and the risk assessment

The items on the risk assessment simplify not only the complexity that emerges from evaluating the ecological context of cases but also the complexity that emerges from debates in the field about what constitutes risk and how risk should be understood. Both the presence of questions related to intimate partner violence and the way these items are scored represent one way of understanding the impact of domestic violence on child welfare. The presence of intimate partner violence is accounted for in two different ways on the risk assessment, making it one of the

most heavily weighted factors to be evaluated during the investigative process. Kate believes that the risk assessment can come out as higher than it should be (in her estimation) simply because of the way domestic violence is understood within the risk assessment:

> Like there's a lot of times like I'm surprised like when it comes out as high. I think that it's not so accurate 'cause there's two different questions regarding domestic violence; one if there's a history; one if the person's in a harmful relationship. So if there's any domestic violence they're scored negatively on two areas, whereas everything else is just one question. So I feel like it's a little skewed when it comes to that. And I don't really know how I feel that if somebody had domestic violence in a relationship five years ago, how that affects their current you know ability to parent and their current relationship if there isn't any domestic violence. I don't necessarily feel like they should be penalized for that you know. So sometimes I think it's a little higher than it should be.

Kate points out that because the risk assessment has two items related to domestic violence *(N8 Primary Caretaker involved in harmful relationships* and A*8 Either caretaker has current or a history of domestic violence*) that families are evaluated twice for the same phenomenon, automatically raising the risk level. As such, the presence of domestic violence, whether in the past or present, could be considered to be one of the areas of greatest concern to Child Protective Services. Adding multiple points to the risk assessment, involvement in a relationship with intimate partner violence might be enough to place a parent on the Central Registry for Abuse and Neglect as Ashley describes:

> A: So now we have referred her to domestic violence resources, but she's out of county now, she's actually moved into Redford. So we're gonna open the case. Unfortunately we have to put her on Central Registry due to the fact that there was neglect going on because she was exposing her child continually to this DV [domestic violence] relationship. This one, I don't feel good about putting her on Central Registry, but we have to open the case to provide a more continuing services.
>
> E: Was there any possibility of putting her at a Cat 3…?
>
> A: We don't make the decision, the computer does, so I put in,

> you know all the stuff that's going on and unfortunately it came out as high risk level so that's that and she's on Central Registry. But we will continue to service the family. We've referred her to getting a PPO. That was our first suggestion and luckily her mom's doors are barred so he can't kick those down so that's actually a big plus. And we're just gonna go from there.

The risk assessment captures one way of thinking about the meaning of domestic violence within Child Protective Services: that exposure to domestic violence is an unqualified risk to children. It is important to note that ample empirical evidence supports this claim (Fantuzzo & Lindquist, 1989; Jaffe, Wolfe, & Wilson, 1990; Kolbo, Blakely& Engleman, 1996; Margolin & Gordis, 2000; Wolak & Finkelhor,1998 (Levendosky, Huth-Bocks, Shapiro, & Semel, 2003). However, the certainty of the risk assessment also obscures an ongoing debate among researchers and practitioners who work on issues related to intimate partner violence about whether charging women with failure to protect in child welfare investigations is itself a process of blaming the victim (Rivett & Kelly, 2006; Rogerson, 2012).

This issue has not reached an empirical or policy consensus, making it far from settled. Workers like Kate and Ashley who take a position that punitive consequences in these cases are unjustified, feel that their hands are tied when caregivers are placed on the Central Registry for Abuse and Neglect or when petitions are filed solely based on the presence of intimate partner violence. While Kate and Ashley's position has support from some researchers and advocates, their understanding of cases involving intimate partner violence goes unexplored because of the way the risk assessment is structured and because state policy does not allow a space for workers to lower the case category.

The way intimate partner violence is captured on the risk assessment may actually serve to create confusion about how to evaluate the role of intimate partner

violence in a case. As Kate explains, official policy for the substantiation of cases is supposed to take a more nuanced view of failure to protect cases that incorporates current debate about the issue:

> And one thing that we always like get caught up on is the sexual abuse cases and whether or not it's you know someone failed to protect or it's a mandatory petition. So those are things that we went around with. We had like three cases in a row, so we were talking about that a lot. And domestic violence is another one. I mean there's a couple pages of very specific things and how you can't substantiate a parent. Like if there's domestic violence, a lot of people don't realize like if Mom's allowing the dad to beat her instead of the children, that's a protecting measure and you can't substantiate Mom for that. She's doing what she can within her means to protect the kids. So you know what I mean? A lot of people interpret that differently because obviously you're still exposing your child to domestic violence, but that's an effort that she's making to protect them.

The static wording of the questions related to intimate partner violence along with confusion among workers about how to evaluate it means that in practice only one side of the debate is recognized in assessments of domestic violence. Issues that are still being negotiated in practice are reified as having reached an expert consensus on the risk assessment itself.

Neglecting Emotional Neglect and Abuse

While some items on the risk assessment are heavily weighted both in terms of number of questions and scoring, others are functionally left out of the assessment process despite their official presence on the risk assessment. One such item is emotional abuse and neglect. This form of maltreatment is explicitly represented on the risk assessment through the following question: *A1: Current complaint and/or finding includes mental injury.* State A's policy regarding mental injury is that only a mental health professional can identify it. Workers report that this requirement means that they rarely, if ever, endorse a finding of mental injury,

answering the question negatively as a matter of routine. Casey and Monica detail this process

> M: Mental injury; No. I don't think I've ever marked that one.
>
> C: And so for the other one mental injury. Mental injury, we don't assess that it's kind of going be assessed by psychiatrist, yeah and so anytime we have an allegation of mental injury we have to send them for an assessment and we wait for their [evaluation]… so I almost always hit zero on this one.

The majority of workers understand this policy to be an unequivocally positive approach, asserting that cases involving emotional neglect or abuse are more likely to be subject to personal judgments rather than objective criteria. Leaving the determination of emotional abuse and neglect to experts protects families from the whims of protective service workers as Casey articulates:

> E: What do you think about sort of assessing emotional abuse and only having a psychiatrist assess that?
>
> C: I don't think that we should be able to assess that…I'm not qualified to make that assessment so I wouldn't think that I should. And I think that we, I don't know that we don't even know some of the things that might, different people or act differently and cope differently to different situations. So things that we might know, like oh, that probably isn't that big a deal; other peoples coping mechanisms may not be as great you know for certain kids and that it is actually mental injury. So I definitely think that we should refer it out to other people.

However, workers also report that they almost never actually seek out a psychiatrist to do this assessment. Assessments of mental injury are omitted from regular consideration during the investigative process even though their inclusion on the risk assessment asks workers to attend to its presence or absence. This means that emotional neglect and abuse are rarely substantiated or addressed by Child Protective Services in State A.

William is one of the few workers who objects to having a mental health professional assess the presence of mental injury because he believes that lack of ability to do so functionally ignores the presence of emotional maltreatment. Below, William describes a case where he disagrees with his supervisor about whether mental injury is occurring:

> Mental injury has to be confirmed by a therapist or a doctor. Okay. Case in particular. Two opposing thoughts; me and my supervisor. A case where the mother, I think as part of discipline to get her kids in line, "I'm sending you back to foster care and I'm never taking care of you again, and you're never gonna see me again." That's mental injury. I also think that that's a poor use of discipline to try to get her kids in line. Right?
>
> [The supervisor says:] 'No, that's not inappropriate discipline and I can't call it mental Injury because it's gotta be substantiated by a therapist, who says this particular thing has really made this child emotionally unstable. All these kids aren't even in therapy for the most part. Why do they have mental injury? Call it something else'.
>
> I think it's mental injury, they're [DHHS administrators are]saying no, you can't use it. And I'm like, when a kid hears I'm going to foster care, I'm going to foster care and has experience in foster care, that maybe wasn't so good what else could it be? Do I really need a therapist to confirm that?

Ultimately unable to convince his supervisor of the presence mental injury in this case, it goes unaccounted for and unaddressed. The policy of permitting only a mental health professional to endorse the presence of mental injury and the subsequent reaction of caseworkers to rarely seek out a professional evaluation represents an incomplete process of quantification for emotional maltreatment. While mental injury remains formally represented on the risk assessment, informally it is still viewed as being too subjective for caseworkers to verify. Left out of the assessment process in practice, mental injury remains able to be quantified only in theory.

Subjectivity within Mechanical Objectivity

Issues related to the best way to handle items that could be subjectively interpreted are a general concern in the construction of actuarial-based risk assessments. *N11. Primary caretaker able to push child needs ahead of own* is one such question whose interpretation remains open and which requirse a different level of skill in scoring than a demographic question such as *A2. Number of prior assigned abuse complaints and/or findings.* Questions that are open to interpretation could lead to exactly what the risk assessment is designed to avoid: disparate assessments of the same case by different workers. In this way, risk assessments may not be as unifying or standardizing in practice as they are considered to be in theory.

In order to manage issues of subjectivity as they relate to scoring the risk assessment, State A has created a policy manual that outlines exactly how questions should be interpreted and what criteria should be used in answering each question. Policy manuals are accessible by clicking on a question when a worker is filling out the risk assessment and through the iPhones that workers carry with them during the workday. Even with these measures, some workers report that they do not ever consult the policy manual, leaving a wide range of responses to the risk assessment open. Workers who do consult the policy manual also recognize that there are individual level differences in how workers respond to questions that can impact whether they are truly standardized instruments. When asking Justine to review how she completed a risk assessment for a particular case, she noted how she struggles to square the complexity of the case with the concreteness of the questions.

> I mean its how you interpret that question you know. Like impulses[Primary caretaker is unable/unwilling to control impulses]; you know my thing here would be she didn't control her impulses to drink, but another worker might not use drinking as, to answer that question. You know it's so I mean yeah like these two questions here…You know like "primary caretaker puts child's needs ahead of her own". That night she didn't but say two days later she's saying all the right things and she's showing, she's willing to put her kids above others. I don't know a worker might say 'okay, she's willing to do what she did but that one instance she didn't so'; it's kind of like depending on how the worker interprets it all.

All risk assessment questions are cross sectional data points, recording information about a caregiver at one moment in time. Workers who attempt to carefully complete the risk assessment can struggle with how to endorse items where the caregiver has behaved differently throughout the investigation. Lila describes the dilemma:

> L: I think you have a lot more like, I think an ability to use your own opinion with the risk assessment versus the safety, that's kind of more cut and dry. But I put no [to whether the caregiver has a substance abuse problem].
>
> E: Okay, you seem embarrassed?
>
> L: Right, well I don't know because I'm like thinking to my supervisor should it be like, you know she tested positive? Like I know, but I tested her again and she tested negative so I don't think she has a problem because if she did she would've tested [positive]…

Lila wants to be as accurate as possible in her assessment but there is no way to record the varied nature of the caregiver's engagement with substances during the course of the investigation. Whether the caregiver has a substance abuse problem is open to some amount of discretion, with some workers using the evidence of a single positive drug screen to confirm and others using the evidence of a single negative drug screen to deny a finding of substance abuse.

The risk assessment encodes a binary understanding of mental health, domestic violence, substance abuse and other phenomena by asking workers to endorse yes/no questions about conditions or states that research has shown to take

place on a continuum and which change over time. Issues that are fluid in nature are encoded linearly and statically. Linking these conditions to a single moment in time, the risk assessment invites interpretation by forcing workers to definitively state the presence or absence of a condition.

Other workers view the risk assessment as entirely open to interpretation and approach it as a way to document their decision-making rather than uncover how cases should be managed. Jennifer does not buy that the risk assessment is a neutral object:

> No, I get that this is numerically based. But, clearly you can manipulate it. And some things, I mean, I've seen people get a high risk level that you wouldn't think; and then I've seen low risk levels where you're like, what? (

Matt explains how manipulating the risk assessment is possible:

> I will change the risk assessment 'cause I will go through it again and say, you know what that might have been a problem, but it's not that big of a problem. Or I will change category 4 to category 3 or category 3 to a 2. I'll adjust it that way.

Workers who understand the risk assessment to capture subjective assessments and who do not view it to be an accurate tool, purposefully fill it out in such a way that it reflects how they would like a case to proceed. For a fuller discussion of how and why workers approach the risk assessment in practice see "Same Difference: Standardized Decision-Making in Child Welfare Practice (Bosk, Same Difference: Standardized Decision-Making in Practice, In Progress). Whether workers consciously or unconsciously use subjective judgments when filling out the risk assessment, the binary phrasing of questions forces workers to make determinative evaluations about phenomena that may be more squishy than concrete. In this way, subjective processes are invited into how the risk assessment is completed. Yet,

these processes remain unacknowledged in the construction of policies related to how the risk assessment is used to determine the trajectory for cases.

Conclusion

Quantifying child welfare decisions through the use of an actuarial-based risk assessment attempts to address a long history of low reliability decision-making in child protective services. Utilizing actuarial-based risk assessments is intended to make decisions both more fair and less biased by basing decisions on empirical research and applying the same standardized criteria to every case. While actuarial-based assessments are considered to be more objective than clinical decision-making, sociological work on commensuration demonstrates that the process of quantification is not entirely a neutral one. Numbers are not completely detached from the social phenomena that organize their use. This paper examines one risk assessment to understand how phenomena related to maltreatment are constituted on the risk assessment and how this organization becomes translated in relationship state policy as cases play out on the ground.

I argue that the organizational and moral imperatives to avoid making a type 2 error may lead to more incidents of type 1 error. The interaction between how risk is constituted on the risk assessment itself and the policies governing its use, may mean that risk is more likely to be overestimated, subjecting families to consequences that are inappropriately punitive as a result. I argue there are two factors that contribute to the likely overestimation of risk: (1) state policy that explicitly ties the score on the risk assessment to a case's trajectory without any formal procedures to lower the risk score and (2) the collection of information only

on the presence or absence of risk factors without a more dynamic understanding of the relationship these factors have to the presenting complaint.

State policy that prohibits workers from lowering a risk score, while building in procedures to raise it, reflects the organizational and moral costs of making a type 1 error. In State A, professional expertise is valued when it raises the risk level but the very real costs related to a risk score that is erroneously lowered leaves little structural incentives to permit clinical judgments that de-escalate cases. Certain types of clinical judgments then are not only seen as less effective than actuarial-based tools but also as dangerous. In practice, policies designed to exclude any clinical judgment from the decision-making process runs counter to the Structured Decision-Making model's original design. The Children's Research Center states:

> The assessment tools are not intended to make case decisions for direct service workers. Staff still need to exercise professional judgment. But the various tools help to structure decisions by bringing objective information to bear on these critical questions(CRC, 2008).

Gambrill and Shlonsky (2001), advocates for the introduction of Evidence-Based Practices in child welfare decision-making acknowledges that

> Use of risk assessment measures may introduce errors, and these should be addressed. For example, workers may rely solely on the instrument and ignore other relevant evidence that the instrument does not address. Family strengths may be ignored if the tool is deficit-based. A tool cannot cover everything and a worker and/or supervisor should be able to override it for good cause (Gambrill & Shlonsky, 2001).

State A's current policies around the risk assessment do not acknowledge that exclusively using the risk assessment to make case decisions might also create the conditions for errors in judgment. Actuarial-based tools contain important information but they are not able to account for all the complexity present in any case. Policies to prohibit caseworker discretion about certain aspects of cases, reflect the power of tropes about quantification, commensuration, and evidence-based

practice at this cultural moment. The very real need for accountability, to reduce the risk of making an incorrect decision, and to address the variability of authority-based decisions facilitates not only the adoption of actuarial-based tools but also policies that limit ecological evaluation of cases. Use of actuarial-based risk assessments without a dynamic evaluation of case context may cause preventable errors as well as reduce them. Such an approach reflects an over determined understanding of both the benefits of actuarial-based risk assessments as well as the weaknesses of clinical decision-making.

Collecting information only on the presence or absence of risk factors without a more dynamic understanding of the relationship these factors have to the presenting complaint creates a structural mechanism by which risk can be overestimated. Phrasing questions in binary terms does not capture the nonlinear nature of many issues that contribute to child maltreatment such as intimate partner violence, mental illness, and substance abuse. Empirical evidence clearly connects all the items on the risk assessment to child maltreatment. Empirical evidence *also* suggests ways in which their risk can be mitigated (e.g. through psychological treatment, social support etc.). Yet, the risk assessment itself does not distinguish between risk factors with and without mitigating circumstances (e.g. a parent that is addressing substance abuse issues vs. one that is not; or a caregiver that was abused as a child and has received treatment for this experience vs. one who has not). It is possible that this more limited conceptualization of risk makes the risk assessment less sensitive in its ability to distinguish between low, moderate and high-risk cases.

A non-relational inclusion of risk factors on the risk assessment reflects the prominence of 'cycle of violence' theories in American culture, which emphasize the

tendency of those who have been exposed to a violent and traumatic behavior to repeat it. From an academic perspective, we can understand the call to 'break the cycle of abuse' to be a shorthand for a set of cultural beliefs about the origins of maltreatment (socially learned) that are then responded to in the form of public policies (in this case tying the score on the risk assessment which uses information about past behavior to predict the likelihood of future maltreatment). While the risk assessment is intended to be probabilistic and not predictive, policies that explicitly tie the risk assessment to case trajectories without mechanisms to adjust their results to account for context conflate the two.

The cultural emphasis on 'the cycle of abuse' can be easily accused of burying the lead. The much larger story is how relatively few children go on to repeat the brutality of their parents. Examining the etiology of child maltreatment, Kaufman and Zigler (1987) call for the childhood myth [of abuse] to be put aside" due to these relatively low documented rates of transmission. A more nuanced understanding of risk and protective factors as they relate to the transmission of abuse is called for. Risk factors that evaluate the relationship between the existence of past abuse and the likelihood for future maltreatment include: the parent's level of traumatic symptoms, level of dissociation, ability to separate from the child, the severity of past trauma, parental representations, parental representations of the child, and fighting-and frightened interactions with the child. Each of these factors have been identified as placing abused children at higher risk for becoming abusive themselves (). Looking at risk factors in a relational way distinguishes between just their presence (which might lead to caregivers inappropriately being placed in higher categories) and their effect on the current charge of maltreatment. While

relational risk factors may be more accurate than the ones currently included on actuarial-based risk assessments, they are also unlikely to be adapted in future iterations. Relational risk factors require an enormous amount of time to detect and a high degree of training to understand. Actuarial-based risk assessments are meant to simplify the evaluation process and revising risk factors so that they are more sensitive may undercut this goal.

In its current form, the risk assessment only collects and evaluates negative information about the family, potentially weighting scores on the risk assessment to overestimate the likelihood of future maltreatment. Protective factors that have been shown to ameliorate the impact of childhood abuse include social support, strong spousal relationships, and positive experiences in therapy. While the Structured Decision-Making model includes an assessment of family strengths through the Family Strengths and Needs Assessment, strengths or experiences that mediate risk factors are not evaluated in relationship to case trajectories. Instead, when cases are opened for services, the Family Strengths and Needs Assessment decides which needs are the three highest priorities to be addressed. Cases presented in this article demonstrate that there are multiple instances where family strengths or mitigating circumstance are not able to be taken into account in case decision-making.

Gambrill & Shlonsky (2001) argue that because actuarial assessments are able to attend to risk in a multiplicative rather than additive way through complex modeling and formulas, they are best positioned to deal with how risk and protective factors interact. I would stipulate that this is only true when risk factors are understood relationally rather than statically. Workers in this article detail the many ways that the risk assessment in State A is ill equipped to understand

complicated interactions between risk factors, protective factors and the larger environment.

Understanding how risk comes to be quantified on a risk assessment and what is privileged and left out as a result is critical to understanding how the risk assessment handles complex and complicated cases. Theory matters because when there is debate about what the meaning of the facts of a case are and how to approach a messy situation, the risk assessment's viewpoint will often settle the debate. In the case where William and his supervisor disagreed about the presence of mental injury, the policy requiring a therapist to indicate its presence functionally meant that William did not determine that mental injury took place. A risk assessment that adds considerable weight to just the presence of intimate partner violence likely influences cases trajectories when there are conflicting opinions about how to proceed.

Throughout case descriptions, workers repeatedly raise concerns about the meaning of policies and instruments that might overestimate case severity and risk. They highlight the serious consequences for families when a caregiver is substantiated for maltreatment at a category that may be inflated. These consequences include job loss, inability to care for grandchildren or relatives at a later time (if their children or relatives are substantiated for maltreatment), and prohibitions from participating in their children's daily lives in the form of field trips. There is a high cost not only to families but also to the child welfare system when families are incorrectly sorted into categories. Intensive services for high-risk families are costly and time consuming. Caseworker time may be inapproriately spent attending to needs that do not require such a high level of services. These

consequences undermine the system as a whole as well as one of the primary goals of the Structured Decision-Making model: accurate classification of cases (Dorsey et. Al., 2008).

Workers reported that approximately 50% of their denied cases score as high risk, raising significant questions not only about the sensitivity of the risk assessment but also the social justice implications of the instrument. If so many denied cases would be treated as some of the most severe in the system if they were substantiated, what does that say about the ability of the risk assessment to distinguish among families? Does the risk assessment simply reflect and refract the disproportionality in the child welfare system that accompanies neighborhood effects or does it contribute to creating these effects when risk in cases are overestimated? Conversely, workers in State A may be more likely to deny cases that should be substantiated because they are working to protect families from facing the most severe consequences of being categorized at a high level. If this is the case, the risk assessment could inadvertently contribute to cases of maltreatment going unaddressed or unacknowledged. More work is needed to examine the relationship between overestimation of risk, disproportionality, state policy and the risk assessment itself.

I make several policy recommendations based on the findings presented above. First, I recommend that workers should be permitted to both raise and lower the risk level of a case. Overrides in either direction incorporate an ecological framework into case decision-making and address concerns by workers that they are not able to account for contextual factors when making case decisions. I do not mean to suggest that workers disagreements with the risk assessment are always correct.

In some cases they may not be. Rather, I suggest that there should be a formal process for resolving issues where caseworker judgments conflict with those on the actuarial-based risk assessment.

I recommend that these formal procedures include requiring workers to make a clinical formulation about the case in relationship to the risk assessment. Clinical formulations would ask workers to integrate and make explicit sense of how the risk factors on the risk assessment impact all the evidence gathered during the investigative process. These formulations could be combined with the formal narrative that caseworkers are already required to submit as part of their administrative tasks. Reflective supervision about cases would help workers to sort out how their own , values, biases, and beliefs influence their interpretations of evidence and their recommendations for a case. A committee composed of supervisors, administrators, and specialists in areas that child welfare cases most often involve (intimate partner violence, mental health, substance abuse, trauma, poverty) could review override requests to ensure that neither overestimation or underestimation of risk occurred as the result of implementing an override.

I recommend adjustments not just to policy in relationship to the risk assessment but also to the risk assessment itself. The risk assessment should be updated to collect information not just on the presence or absence of risk factors but also on their relationship to the current complaint. While the most sensitive risk factors might be difficult to collect as child protective services work is currently organized, the phrasing of questions could be altered to be more dynamic. For example, questions on the presence of substance abuse could be reframed from binary to reflective of a continuum. Some possibilities for revisions include: (1)

Primary caretaker has had a positive substance abuse screening during the course of the investigation (with points given to reflect the number of positive screenings) (2) *Primary caretaker is currently in treatment to address substance abuse treatment. (2b) If Primary caretaker is not currently in treatment to address substance abuse, primary caretaker is willing to participate in services.*

Demographic questions could also be reformulated to contextualize the presence or absence of documented risk factors within the current case. For example, questions about the number of children in the house could be revised as follows: *There are more than three children in the house Y/N. If Y, Primary caretaker is overwhelmed by meeting the needs of all the children.* Phrasing questions in this way has the potential to make the presence or absence of risk factors directly relate to the current complaint and the assessment process. A larger issue for the field is how to include the risk factors that require more expert judgment such as those related to parental mental representations and parent-child relationships in the risk assessment process. Just because this would be difficult to do does not mean that it should not be done. What is the purpose of actuarial-based risk assessments if they do not collect the most accurate information or information that is most relevant to determining the likelihood of future risk of maltreatment?

The State A risk assessment has not been revised since 2008. If the goals of risk assessments are to link child welfare decision-making to the latest empirical research then regular revision, updates, and reviews to the risk assessment itself are required. My final recommendation is that the risk assessment be revised every 5-10 years to reflect the latest findings in the field and revisit theoretical debates in the field. Part of the work of revising the risk assessment is not just updating

statistical models to integrate new research and understanding but also a deeper assessment of the epistemological foundations of the actuarial-based risk assessment in use in relationship to child protective service work. During periods of revision, explicitly talking with workers about how the risk assessment informs their work, when they disagree with the risk assessment itself, and how they are utilizing it in practice has the potential to reveal how processes of commensuration are working as intended and unintended.

In this paper, I have demonstrated that the development and use of actuarial-based risk assessments are part of a process of quantifying child welfare decision-making. Quantifying decision-making in child welfare embeds specific values within the risk assessment process itself. While the State A risk assessment attempts to make child welfare decisions more fair, less biased, and more accurate it also may have the unintended consequence of overestimating risk for certain populations based on how the questions are phrased and interpreted on the risk assessment itself in combination with the policies related to its use.

Footnotes

[1] Heading taken from title of Espeland and Stevens Annual Review of Sociology article by the same name.

[2] In State A, the Structured Decision-Making model is compromised of a series of actuarial-based risk assessments including: (1) an initial intake assessment that decides which calls will be investigated; (2) a safety assessment that determines whether it is safe for a child to remain in the home during the course of the investigation; (3) the risk assessment which dictates the trajectory of a substantiated case; (4) the family strengths and needs assessment and (5) the child strength and needs assessment which prioritize the services families and children who are being serviced by the child welfare system will receive; (6) a risk reassessment with considers whether the risk level for an active case has changed; and (7) the reunification assessment which analyzes whether children who have been removed should return home.

Works Cited

Bowker, G., & Starr, S. (1999). *Sorting Things Out*. Cambridge, MA: MIT Press.

Bowlby, J. (1973). *Attachment and Loss: Separation: Anxiety and Anger* (Vol. 2). New York: Basic Books.

Bowlby, J. (1979). *The making and breaking of affectional bonds*. New York: Basic Books.

Bowlby, J. (1980). *Attachment and Loss: Loss Sadness and Depression* (Vol. 3). New York: Basic Books.

Coulton, C. J., Korbin, J. E., & Su, M. (1999). Neighborhoods and Child Maltreatment: A Multi-Level Study. *Child Abuse & Neglect, 23*(11), 1019-1040. doi: http://dx.doi.org/10.1016/S0145-2134(99)00076-9

Coulton, C. J., Korbin, J. E., Su, M., & Chow, J. (1995). Community Level Factors and Child Maltreatment Rates. *Child Development, 66*(5), 1262-1276.

CRC. (1999). The Improvement of Child Protective Servies with Structured Decision Making: The CRC Model. In NCCD (Ed.).

CRC. (2008). The Structured Decision-Making Model: An Evidenced-Based Approach to Human Services: Children's Research Center-A Division of the National Council on Crime and Delinquency.

Epstein, S. (2007). *Inclusion: The Politics of Difference in Medical Research*. Chicago: The University of Chicago Press.

Espeland, W. (1997). Authority by the Numbers: Porter on Quantification, Discretion, and the Legitimation of Expertise. *Law and Social Inquiry, 22*.

Espeland, W., & Sauder, M. (2007). Rankings and Reactivity: How Public Measures Recreate Social Worlds. *American Journal of Sociology, 113*(1), 1-40.

Espeland, W., & Stevens, M. (1998). Commensuration as a Social Process. *Annual Review of Social Science, 24*.

Espeland, W., & Vannebo, B. (2007a). Accountability, Quantification, and Law. *Annual Review of Law and Social Science, 3*.

Espeland, W., & Vannebo, B. (2007b). Accountability, Quantification, and Law. *Annual Review of Law and Social Science, 3*, 21-43. doi: 10.1146/annurev.lawsocsci.2.081805.105908

Fluke, J. D., Yuan, Y. Y. T., Hedderson, J., & Curtis, P. A. (2003). Disproportionate Representation of Race and Ethnicity in Child Maltreatment: Investigation and Victimization. *Children and Youth Services Review, 25*(5-6), 359-373. doi: 10.1016/s0190-7409(03)00026-4

Gainsborough, J. (2010). *Scandalous Politics: Child Welfare Policy in the States*. Washington DC: Georgetown University Press.

Gambrill, E., & Shlonsky, A. (2000). risk assessment in Context. *Children and Youth Services Review, 22*(11-12), 813-837.

Gambrill, E., & Shlonsky, A. (2001). The Need for Comprehensive Risk Management Systems in Child Welfare. *Children and Youth Services Review, 23*(1), 79-107.

Garbarino, J., & Crouter, A. (1978). Defining the Comminity Context for Parent-Child Relations: The Correlates of Child Maltreatment. *Child Development, 49*(3), 604-616.

Garbarino, J., & Kostelny, K. (1992). Child Maltreatment as a Community Problem. *Child Abuse and Neglect, 16*(4), 455-464. doi: 10.1016/0145-2134(92)90062-v
Garbarino, J., & Sherman, D. (1980). High-Risk Neighborhoods and High-Risk Families: The Human Ecology of Child Maltreatment. *Child Development, 51*(1), 188-198.
Hacking, I. (1990). *The Taming of Chance*. Cambridge: Cambridge University Press.
Lampland, M., & Starr, S. (1999). *Standards and Their Stories: How Quantifying, Classifyign and Formalizing Practices Shape Everyday Life*. Ithaca, NY: Cornell University Press.
Latour, B., & Woolgar, S. (1986). *Labratory Life: The Construction of Scientific Facts*. Princeton: Princeton University Press.
Levendosky, A. A., Huth-Bocks, A. C., Shapiro, D. L., & Semel, M. A. (2003). The Impact of Domestic Violence on the Maternal - Child Relationship and Preschool-Age Children's Functioning. *Journal of Family Psychology, 17*(3), 275-287.
Porter, T. (1995). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
Power, M. (1997). *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.
Radin, M. (1996). *Contested Commodoties*. Cambridge, MA: Harvard University Press.
Rivett, M., & Kelly, S. (2006). 'From awareness to practice': children, domestic violence and child welfare. *Child Abuse Review, 15*(4), 224-242. doi: 10.1002/car.945
Rogerson, S. (2012). Unintended and unavoidable: The failure to protect rule and its consequences for undocumented parents and their children. *Family Court Review, 50*(4), 580-593. doi: 10.1111/j.1744-1617.2012.01477.x
Timmermans, S., & Epstein, S. (2010). A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization. *Annual Review of Sociology, 36*.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainity: hueristics and biases. *Science, 185*, 1124-1131.
Tversky, A., & Kahneman, D. (2002). Extensional vs. intuitive reasoning: the conjunction fallacy in probabilistic judgment. In T. Gilovich, D. Griffin & D. Kahneman (Eds.), *Hueristics and biases: the psychology of intuitive judgment*. New York: Cambridge University Press.
Zelizer, V. (1985). *Pricing the Priceless Child*. New York: Basic Books.
Zuravin, S. J. (1989). The ecology of child abuse and neglect: Review of the literature and presentation of data. *Violence and Victims, 4*(2), 101-120.

Chapter 5

Conclusion

This dissertation explores how actuarial-based risk assessments are implemented to facilitate child welfare decision-making. I examine the intended and unintended consequences of locating decision-making about child welfare in actuarial-based tools at three different levels of analysis: individual worker, agency, and the risk assessment itself. Despite the widespread diffusion of actuarial-based assessments to frame child welfare decision-making, we have little understanding of how these technologies are implemented on the ground and under what conditions differences in organizations, individuals, and models influence their use. Limited work in this area suggests that workers ignore actuarial-based risk assessments in favor of clinical judgment or engage in actuarial-based decision-making as a strategy for negotiating bureaucratic policies and rules (Hilton & Simmons, 2001; Krysik & LeCroy, 2002; Lyle & Graham, 2000).

While questions about the role risk assessments have in informing decision-making can seem dryly theoretical in nature, the consequences for how we create, organize, and implement policies related to child welfare decision-making are anything but academic to the children the child welfare system is intended to protect and the families it is charged to serve.

Using a three article format, I asked the following questions: 1) How do

structured decision-making aids influence decision-making in the context of every day work routines? (2) How do standardized tools function across practice settings? And (3) What are the consequences for how one state's risk assessment currently quantifies risk? Grounding my analysis of the empirical data within sociological theories of power and privilege, commensuration, and organizational theory, I understand the construction, experience, and every day use of the actuarial-based risk assessment to be both a technical and social process. The meaning(s) and expression of the actuarial-based risk assessment take form dialectically. State and organizational policy dictate how the tool should be understood, deployed, and utilized in every day work, which in turn, shapes and is shaped by, individual responses and informal work practices.

This research seeks to contribute to sociological and social work theory by shifting discussion about the use of clinical judgment and evidence-based practice from a dichotomous framework to an integrative one. Interrogating the polarization of these concepts as they play out on the ground reveals a false split. In reality, when organizations and individuals use the actuarial-based risk assessment they are acting from both models simultaneously, complicating discussions about the utility and drawbacks of employing empirical or clinical approaches in an exclusive way. My work attempts to trace out how workers and organizations contend with the messiness of competing paradigms. These paradigms are framed within policy as separate from one another but are intersecting, complementary, and competing on the ground. Each paper examines how this is done differently:

In Chapter 2, I explore how workers contend with Structured Decision-Making model within a tightly controlled decision-making environment. Rather than having

a uniform response to the Structured Decision-Making model, the extent to which workers follow the model varies widely. I find that power and privilege influences which workers feel that they can flout formal policy and rules and which workers feel constrained to follow policy even when they strenuously disagree with it. How workers conceptualize the actuarial-based risk assessment also informs how they apply it to their casework. When workers understand the risk assessment to be unsuccessful in achieving a standardizing function, they feel that they are not doing anything differently than their colleagues when they explicitly adjust their formal responses to questions to get the outcome they desire. Conversely, workers who believe the risk assessment to be successful in constraining individual subjectivity and bias use it enthusiastically and unquestioningly. Finally, I observe that all workers utilize discretion at some point in their casework, regardless of how they conceptualize the risk assessment. Workers report that they use the substantiation process to exercise their professional authority and judgment. In this way, discretion does not disappear from the process, it just changes form (Espeland & Vannebo, 2007).

Chapter 3 examines how the Structured Decision-Making model is implemented in two different states. Comparing the use of the risk assessment in a tightly controlled decision-making environment to a loosely controlled one, I contend that the organizational context in which it is utilized shapes the meaning of the tool itself. By demonstrating that State B policy creates a loose connection between the risk assessment and case trajectories, making it functionally meaningless and that State A policy creates a tight connection between the risk assessment and case trajectories that often excludes environmental context into case decision-making, I

highlight the difficulties with both implementation strategies. These case studies offer empirical examples of how evidence-based practices are not self-actuating and complicate the discussion of how to adapt a tool to the local context.

In Chapter 4, I look at how the State A risk assessment itself structures risk. Policies that emphasize avoiding errors where a child is mistakenly left in the home may then underemphasize the serious implications for inaccurately classifying families as high risk. Without formal procedures to de-escalate cases, workers report that families can be treated punitively based on having the wrong mix of demographic factors present in a case.

Key Themes and Implications for Social Work Research and Practice

Considering these chapters together provides the opportunity to draw out and synthesis the key themes that run across them. These themes command sustained discussion in their own right and this next section analyzes what we learn from thinking about this work as a whole.

One of the analytic strengths of this research lies in comparing how policies and practices related to the Structured Decision-Making Model differ across states. The striking contrasts in the weight the SDM is given in decision-making in State A and State B highlight the need for contextualizing all discussions about evidence-based practice. **How an intervention works, whether it works and why it works are as much a function of the social context of the intervention as the intervention itself.** The differences between State A and State B flow from how use of the SDM is connected to organizational policy and how this policy became translated into practice. The incredibly local nature of child welfare policy makes it unlikely that

any intervention intended to unify child welfare practice across settings will be able to do so. Researchers and practitioners need to expect that interventions will change when they are deployed and that these changes may differ from setting to setting. Instead of framing the multiplicity of factors that influence the meaning of an intervention in a particular context as something to always be controlled and managed, a starting point for implementation science should be to develop theory about how to begin to understand interventions in relationship to the settings in which they will be deployed and what the consequences of this relationship may be. Dealing with the messy and ephemeral parts of translation offers opportunities to re-conceptualize how we think about interventions themselves. The Practice-Based Evidence Movement and research on the common factors of interventions is moving in this direction but has not yet fleshed out what this looks like in practice (Fox, 2003; Okpych & Yu, 2014).

The standardizing function of the SDM is related to both organizational imperatives and individual approaches. The SDM is designed to limit the role that social identity, social context, and organizational policy play in case outcomes. My findings show that in a loosely implemented setting (State B) that the SDM may indeed have standardized how workers approached the risk assessment but not in the way intended. Because the risk assessment did not carry much weight in the final outcome of the decision, workers did not need to develop strategies about what to do when they disagreed with its result, they just simply ignored it. The less the SDM was emphasized, the more likely it was to offer an assessment of the case without conscious adjustment of the worker (ironically standardizing how workers utilized it). However, this standardization becomes meaningless because the very

thing that allows workers to fill the risk assessment out freely (its lack of formal decision-making power) is also the thing that allows workers to ignore the risk assessment in their decision-making.

Conversely, the more power policy gave to the SDM in decision-making, the more individual expressions of authority began to operate in how some workers approached the SDM and the less standardized worker's approach to it became. The effectiveness of policy in tying case outcomes to the SDM opened the door to some workers finding other ways to influence case outcomes.

Its not whether there is discretion in standardized processes but how, in what form, and with what consequences. Discretion was practiced differently by different workers and organized and controlled differently by different agencies. Discretion was present in the most tightly controlled implementation of the SDM and some elements of standardization were present in the loosest. While workers might not have used much discretion in the way they filled out the SDM in State B, they did use an enormous amount of discretion in how they approached and arbitrated their cases. Despite the presences of variation in how workers approached the SDM in State A, there was less opportunity to exercise discretion in case decision-making overall and expressions of discretion that ran counter to policy were limited to a select population of workers.

These results highlight that discretion operates on a continuum across and within settings. This fact is rarely reflected either in social work theory or in practice. Instead, discretion is often presented, debated, and examined as an all or nothing practice. Our theory badly needs to catch up with the differential ways that discretion operates organizationally and individually within attempts to constrain it.

The continual presence of discretion in varied forms represents one of the most robust findings from this research that is likely generalizable across fields such as medicine, education, and law where actuarial-based assessments are used to define work practices. Understanding where and when discretion is utilized would help untangle how expert judgment and actuarial-based or empirical technologies work alongside each other. Further, this work highlights that even within an organization, responses to an intervention are not homogenous but heterogeneous. Research needs to be more sensitive to this fact.

Power and privilege may become reproduced in how discretion is expressed and by whom it is performed, when policies are developed to constrain it. This research raises questions about how issues of power and privilege become produced and reproduced within contexts intended to limit individual judgment. Because the practice of discretion is so deeply connected to the performance of authority, it makes sense that when discretion is constrained, that people who are used to having a voice will work to find a way to have one again. Because the expectation of having an individual voice is often related to being in a position of social privilege, finding ways to maintain discretion may reproduce existing power relations in aggregate. Throughout my research, those who actively resisted organizational attempts to limit their subjective readings of the facts of the case came up with successful strategies to influence case trajectories. Conversely, those who were not used to having a voice did not expect this to be a part of their work, no matter how negatively they felt about the processes in which they were engaging. People who actively resisted constraining policies tended to have a combination of social status characteristics associated with privilege, suggesting

that even within policies meant to flatten social difference, that power differentials can and do operate in ways that may not be noticed by individuals themselves or the larger system as a whole.

At the same time, systems that openly leave a space for individual discretion may not avoid reproducing power and privilege differentials that emerge from peoples' social locations either. In fact, we have lots of research from many different fields to suggest that it does not. It is not clear from the data presented here, in what way, if any, power and privilege influenced how individuals approached their work in State B. The focus of my study was not to investigate how these concepts operated so I was not able to collect data that might have revealed them. However, this does not mean that powerful social inequalities were not reproduced in other ways in the organizational environment and/or with regards to case decision-making. I would like to leave open the possibility that power and privilege influenced how decisions were made and what those decisions were in State B in ways that this work was not able to detect.

Power and privilege are central organizing concepts in social life. They are present in almost every social situation and work implicitly and explicitly to shape individual and collective experiences and approaches. The SDM represents one attempt to even the playing field for families and for workers. It is also a profoundly logical one. By raising the ways in which power and privilege come into play in people's response to the SDM, I do not mean to suggest that the SDM is completely ineffective at constraining them. Instead, I hope to highlight the ways in which power and privilege emerge in unexpected ways. I think this speaks to the force of power and privilege as organizing social concepts. Across interventions, social work

researchers must be alert to the ways in which social inequalities and more subtle expressions of social difference play out. In medicine, law, social work, education, we should expect and assume that power and privilege will be in effect in some capacity even as we work to limit their impact. Theoretically, empirically, and practically, we must remain conscious and curious about the ways in which power and privilege become produced and reproduced within the interventions and policies meant to flatten them.

Practices intended to simplify make other things complicated. As with all processes of commensuration, the use of actuarial-based assessments in child welfare decision-making is partially intended to simplify an extremely complex process. While the actuarial-based risk assessment simplifies some aspects of decision-making (e.g. How to think about the meaning of the presence of specific risk factors related to future maltreatment), this process also complicates others. Choices organizations make about how the tool will be deployed combined with decisions about how to constitute the risk assessment itself shape what parts of decision-making become simplified and what become complicated.

For example, State A policy requiring a mental health professional to verify the presence of mental injury is intended to simplify the assessment process for emotional maltreatment. Instead of introducing the difficulties that accompany having Child Protective Services (CPS) workers make a judgment about such a subjective phenomena, policy has been developed to streamline the process. Functionally, this process of simplification complicates the ability for Child Protective Services) in State A to detect or deal with emotional maltreatment when it does occur. Workers almost uniformly reported that the trouble (in all senses of

the world) related to getting a mental health professional to verify this diagnosis means that they essentially ignore the question. As a result, emotional maltreatment may not be effectively counted, assessed, or intervened with when it is a part of child welfare cases.

What becomes complicated through processes of simplification is often obscured. As social work researchers develop, deploy, and evaluate interventions, attention must be paid to not just to whether an intervention is working but also to how the intervention has changed work processes and what these changes have meant for how problems or issues are understood and managed. We should expect that simplifying some things is likely to have the unintended consequence of complicating others.

Standardized processes do not allow for informal solutions to problems. The Structured Decision-Making model is designed to unify how work is done and what the outcomes of that work are among Child Protective Service workers. The formal rules and regulations related to ensuring that all complaints are handled in the same way serve multiple purposes: to provide mechanisms for accountability, to create efficiency, to develop objective criteria for decisions, ground decisions in empirical evidence, reduce error, fill gaps in the training and skill of workers, and ensure that blatant bias does not enter into the decision-making process. At the same time these formal rules and procedures prohibit workers from managing issues in ways that are counter to the way procedure dictates but nevertheless may be appropriate.

Workers in State A report that at least half their caseload involves investigating bogus complaints that are vengeful in nature. Investigators told me about cases

where they were receiving weekly reports about a family, which they knew to be unfounded. Standardized procedures prohibited investigators from handling these cases any differently than they would other complaints. Yet, without the ability to either forgo some aspects of the investigation or come up with another way to approach the case, CPS workers must proceed in a manner that does not make much sense. In these cases, following procedure may cause unintended harm when children are repeatedly asked questions they have already answered and the (unlawful) behavior of the reporting person is continually reinforced. Cases involving multi-generational families that must be separated due to a complaint against one of the caregivers but not any of the others, represent another area where workers believe they could solve the issue more efficiently, more fairly, and with less cost to the system then policy allows. A consequence of standardizing procedures then is the inability to respond logically to situations that are not anticipated in formal procedures but nonetheless emerge and must be managed.

One way that actuarial-based assessments function are to clarify the parameters for decisions. As a result, actuarial-based tools reify singular understandings of particular issues. **In doing so, they settle issues that are unsettled in theory and in practice.** Workers reported confusion and disagreement over how cases involving intimate partner violence are captured and assessed on the risk assessment in State A. This confusion mirrors on-going debates in the field about how to think about, and what to do about, children's exposure to child maltreatment. The empirical literature is clear that exposure to intimate partner violence can have negative socio-emotional impacts on children (Fantuzzo & Lindquist, 1989; Jaffe, Wolfe, & Wilson, 1990; Kolbo, Blakely& Engleman, 1996;

Levendosky, Huth-Bocks, Shapiro, & Semel, 2003; Margolin & Gordis, 2000; Wolak & Finkelhor,1998). However, advocates point out that women may have little choice but to remain in violent relationships and that it is unjust to penalize women with involvement in the child welfare system (Rivett & Kelly, 2006; Rogerson, 2012).

. Even though State A policy takes a more expansive view on the decision to substantiate for exposure to child maltreatment, the risk assessment itself weights exposure to intimate partner violence heavily. Cases that receive a substantiation where intimate partner violence is present are likely to be quickly escalated through the child welfare system. In this way, the risk assessment obscures practice debates and State A policy offers no procedures for resolving disagreements when the risk assessment considers a case to be more severe than the worker. Current policy does not acknowledge that the risk assessment may capture a definitive view of an in issue that remains unresolved as a whole. New procedures must be developed to address such moments.

Without regular updating of actuarial-based assessments, outdated social work approaches or major shifts in the field may not be captured for long periods of time. As knowledge is created, as new debates emerge, as practice changes, so to do our understandings of what creates risk and how risk should be evaluated. Of course, formal design of the Structured Decision-Making Model includes regularly revision to the model itself to capture new information in the field. It is one thing to say the SDM should be updated regularly and another to do it. Going through the process of revision involves managing bureaucratic requirements, funding an expensive task or simply making time for an intensive process in a crisis driven field. In short, there are significant barriers to regularly

revising the risk assessment. This raises questions about how to structure or require formal revisions or updates in the field. Is it realistic to expect that states will invest the time to do so? If not, how much weight can be given to the risk assessment and for how long? These are open questions that need to be carefully answered.

It is also important to acknowledge that change presents costs to the system as well. Workers in State B shared that the regular changes in policies and procedures made it hard to take any new system or idea seriously. Many workers engaged a strategy of just 'waiting it out' when they did not want to adopt a new approach to their work with the confidence that the rules would change soon enough. As a whole, researchers must think about the time frame for continual updating of interventions. What is the happy medium between change that is not too much or too little?

Finally, this research calls attention to the fact that **resistance to evidence-based practice, at times, is grounded in real critiques of the intervention.** The implementation science literature often understands resistance to the adoption of interventions to be a flaw in individual workers that must be fixed. The prevailing wisdom goes that with the right amount of education workers will understand the necessity for Evidence-Based Practice and will make the appropriate changes to their practice. This understanding of resistance to Evidence-Based Practices overlooks the possibility that refusal to accept new interventions may be grounded in very real critiques of them. Workers critiques of the SDM in both states revealed serious tensions between the theory and implementation of the SDM, tensions that were likely unanticipated by those not in the field. As we evaluate interventions or consider the ways in which implementation falls short of our expectations, we must

invite the voices of those on the front lines into the discussion. What can we learn about the relationship between theory and practice from those that struggle with adapting an intervention from the ideal to the real?

Joining the Clinical with the Empirical: A Both/And Not an Either/Or

The data presented throughout the dissertation gives life to Timmerman's and Epstein's assertion that "on the ground, every standard is simultaneously over determined and incomplete"(Timmermans & Epstein, 2010)." Without flexible ways to respond to the SDM and the policies structuring its use, workers are left with several choices when encountering a situation for which policy is over determined: (1) bend policy to address a situation that is not covered in formal procedures; (2) endorse a decision which they feel is not supported by the facts of the case or with which they disagree; (3) deny a case that should be substantiated out of concern for how the case would necessarily unfold. When the nuances of a case do not align with policy, we can understand that policy to be incomplete as any one rule cannot possibly take into account all the contingencies and complications that characterize a messy world. When workers have a difficult time making sense of the evidence before them or the case itself, it may be because they do not have a clear way to integrate the complexities of a case with policy as it is written.

How are we to reconcile and address the ways in which all standards are simultaneously over determined and incomplete? I believe one way this can be done is by joining technologies to standardize clinical judgment with reflective practices. Throughout this work, I have shown that clinical judgment already exists alongside and within procedures to constrain discretion. Given this fact, it makes sense to formalize the integration of discretion into standardized processes. Doing so would

offer a way for the exercise of discretion to support the goals of Evidence-Based Practice rather than subvert them. Analyzing the strengths and drawbacks of the Structured Decision-Making Model's implementation in both states reveals that for interventions to be maximally successful, they need to combine opportunities to exercise expert judgment and express disagreement within standardized processes for decision-making. Building in spaces for the expression of expert judgment and dissent create a formal mechanism by which organizations and individual workers can resolve moments in which policies are over determined and/or incomplete. I have recommended several ways for this integration of reflective and standardize processes to take place. In summary they are:

Require a narrative component to accompany the risk assessment or any other parts of the SDM that are foundational for decision-making. Including a narrative portion of the risk assessment offers a space for workers to formally agree or contest the final score *in relationship* to the empirical evidence considered on the risk assessment. Through a narrative, workers would have the space to offer alternate explanations and highlight the ways in which the nuances of the case counter the finding of the risk assessment or, alternatively, endorse the results. This process would engage all workers at all levels of skill. It would formalize a space for workers who expect to have a voice to offer their opinions and it would invite workers to offer their insights who do not feel entitled to do so.

Literature on procedural fairness in psychology and law illustrates the importance of incorporating reflective processes into decision-making. Porcedural fairness can be glibly summed up as: when people feel the procedure is fair, they are likely to feel good (or at least better) about the result, even when the outcome does

not go their way (Folger, 1977; Sedikides, Hart, & De Cremer, 2008; van den Bos, Vermunt, & Wilke, 1997). Rigid implementation of actuarial-based risk assessments (which runs counter to the original design) does not create a space for people's views to be heard. A narrative component of the risk assessment or the SDM would provide a mechanism for procedural fairness to take place by giving workers a chance to confirm or challenge the official narrative of a case with their own view. This does not mean that workers assessments would override the ones on the risk assessment, only that there would be a formal place to capture a workers understanding of the case. Integrating narrative and actuarial processes might improve decision-making across contexts. In places with a loose implementation like State B, where child welfare workers do not regularly take the score on the risk assessment into account when making case decisions, requiring a narrative component would force them to do contend with this information and apply it to their recommendation. Finally, including narrative components on the risk assessment also offers a way to create narrative coherence between the official narrative of a case and the workers personal one. Even if these narratives differ, the worker herself would have a formal space to wrestle with the ways in which they clash and in doing so create a way for these narratives to exist together in the official record.

Formal procedures should be designed to allow workers to escalate and de-escalate cases. Formal procedures that would allow workers to escalate and de-escalate cases would enhance their ability to take into account all the ecological factors present in cases. Informal solutions to problems could be more easily pursued if workers are allowed overrides in both directions. A careful process would

need to be constructed to manage the risk that de-escalating a case entails. I suggest that a committee composed of supervisors, administrators, and specialists in areas that child welfare cases most often involve (intimate partner violence, mental health, substance abuse, trauma, poverty) review override requests to ensure that neither overestimation nor underestimation of risk occurs as the result of implementing an override. The override process itself would call for formal reflection about the ways in which the actuarial-based assessment understands the case and why this understanding differs from the worker's or organization's perspective.

Limitations of the Study

This research is limited in several important ways. First, as a qualitative study built upon an interpretivist epistemology, its results are not generalizable beyond the sample reported. The diffuse nature of child welfare policy makes it difficult to extend conclusions beyond the states discussed here. Instead, these results are suggestive of avenues for future research, raise complications to prevailing wisdom, and empirically flesh out issues that so far have only been contended with at a theoretical level.

My positionality as a young, white woman, pursuing my Ph.D. may have influenced the responses to questions in ways that I am unaware. It is possible that workers in State A that I typed as restricted in their decision-making where actually more autonomous but did not feel free to tell me about the ways in which they altered their judgments on risk assessments because of a fear of judgment or retaliation. Conversely, it is also possible that workers I typed as autonomous engaged in a form of braggadocio that was unreflective of their actual practices when they told me that they manipulated the risk assessment to reflect a judgment

that they had already made. In State B, I was introduced to study participants through agency directors. This official tie may have made some people reluctant to tell me the full story of how they approached their work if they feared that my promises of confidentiality would not be fully honored. I have no way of knowing if any of these limitations were issues.

Throughout this work, I investigate the ways in which power and privilege play out in contexts were policies and procedures are designed to limit their expression. These concepts require investigation in other settings, using much larger samples, to see if these findings can be applied more broadly.

Although I hope that this has been clear throughout this dissertation, one point bears repeating here: This work is not an evaluation of whether the SDM works nor is it a study of its effectiveness. I make no claims about the overall validation of the SDM in State A and State B. By restricting my analysis to how the SDM is used in practice, I hope to fill a gap in our understanding of what happens when this model is used on the ground to frame child welfare decision-making.

Directions for Future Research

The work that I have presented here suggests several avenues for further inquiry. I have suggested that processes of reflective functioning be integrated with standardized assessments to explicitly incorporate the hidden ways that discretion can operate in decision-making processes. Empirical work is needed to develop and test interventions using this model to evaluate whether such integration makes sense. Just as I have asked here: What are the intended and unintended consequences of pursuing this strategy? What gets left in and what gets left out of decision-making that uses an integrated model to facilitate child welfare decision-

making? What becomes simplified and what becomes complicated as a result?

Research in the field of infant mental health highlights how processes of mentalization and reflective functioning are key to negotiating trauma and improving parent-child relationships (Fonagy, Steele, & Steele, 1991; Fonagy, Steele, Steele, Moran, & Higgit, 1991; Lieberman, Van Horn, & Ippen, 2005; Slade, 2005). In this work, I have extended these concepts to apply to child welfare decision-making and the utilization of actuarial-based risk assessments. At the moment, discussions about mentalization and reflective functioning are restricted to the individual or to dyadic relationships. However, I believe that there could be great power in applying these ideas to the level of systems. What would it look like to have systems like child welfare be more reflective about their processes as whole? How could we move these discussions up a level theoretically and empirically? Why not use some of the most promising interventions in the field to the systems that struggle with some of the same problems as families? These are questions that I have only begun to grapple with but this research indicates that there are many unexplored possibilities here.

Finally, this work suggests that relational and contextual concepts need attention within actuarial-based assessments. Traumatic events are currently constituted on State A's risk assessment in a flat way that workers report often penalizes the person who experienced them. This view reflects an empirical understanding of the lasting psychological impacts of trauma and its multigenerational effects. How might assessments of trauma and mental health be captured on risk assessments in particular and/or in the child welfare system in general that honor their very real impact but also does not pathologize or penalize?

What would a trauma informed child welfare risk assessment look like? How can the organizational imperative of child welfare agencies to mitigate risk be balanced with the social imperative to help families that are struggling? Answering these questions are essential to having a fair and just child welfare system.

Works Cited

Espeland, W., & Vannebo, B. (2007). Accountability, Quantification, and Law. *Annual Review of Law and Social Science*, *3*, 21-43. doi: 10.1146/annurev.lawsocsci.2.081805.105908

Folger, R. (1977). Distributive and Procedural Justice: Combined Impact of Voice and Improvement on Experienced Inequity. *Journal of Personality and Social Psychology*, *35*(2), 108-119. doi: 10.1037/0022-3514.35.2.108

Fonagy, P., Steele, H., & Steele, M. (1991). Maternal Representations of Attachment during Pregnancy Predict the Organization of Infant-Mother Attachment at One Year of Age. *Child Development*, *62*(5), 891-905.

Fonagy, P., Steele, M., Steele, H., Moran, G., & Higgit, A. (1991). The Capacity for Understanding Mental States: The Reflective Self in Parent and Child and Its Significane for Security of Attachment. *Infant Mental Health Journal*, *12*(2), 201.

Fox, N. J. (2003). Practice-Based Evidence: Towards Collaborative and Transgressive research. *Sociology-the Journal of the British Sociological Association*, *37*(1), 81-102. doi: 10.1177/0038038503037001388

Hilton, N. Z., & Simmons, J. L. (2001). The Influence of Actuarial Risk Assessment in Clinical Judgments and Tribunal Decisions about Mentally Disordered Offenders in Maximum Security. *Law and Human Behavior*, *25*(4), 393-408.

Krysik, J., & LeCroy, C. W. (2002). The Empirical Validation of an Instrument to Predict Risk of Recidivism among Juvenile Offenders. *Research on Social Work Practice*(12), 71-81.

Lieberman, A. F., Van Horn, P., & Ippen, C. G. (2005). Toward Evidence-Based Treatment: Child-Parent Psychotherapy with Preschoolers Exposed to Marital Violence. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*(12), 1241-1248.

Lyle, C. G., & Graham, E. (2000). Looks can be Decieving: Using a Risk Assessment Instrument to Evaluate the Outcomes of Child Protection Services. *Children and Youth Services Review*, *22*, 935-949.

Okpych, N. J., & Yu, J. L. H. (2014). A Historical Analysis of Evidence-Based Practice in Social Work: The Unfinished Journey toward an Empirically Grounded Profession. *Social Service Review*, *88*(1), 3-58. doi: 10.1086/674969

Sedikides, C., Hart, C. M., & De Cremer, D. (2008). The Self in Procedural Fairness. *Social and Personality Psychology Compass*, *2*(6), 2107-2124. doi: 10.1111/j.1751-9004.2008.00156.x

Slade, A. (2005). Parental Reflective Functioning: An Introduction. *Attachment and Human Development*, *7*(3). doi: http://dx.doi.org/10.1080/14616730500245906

Timmermans, S., & Epstein, S. (2010). A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization. *Annual Review of Sociology*, *36*.

van den Bos, K., Vermunt, R. l., & Wilke, H. A. M. (1997). Procedural and Distributive Justice: What is Fair Depends More on What Comes First than

on What Comes Next. *Journal of Personality and Social Psychology 72*(1), 95-104. doi: 10.1037/0022-3514.72.1.95