Hands On Testing (HOT) Pilot Project

Recommendations and Findings: Final Report

Presented to the Oklahoma Department of Human Services, Child Welfare Services

Julie Miller-Cribbs, Kirsten Havig, David McLeod, Linda Smith

The UNIVERSITY of OKLAHOMA
Anne and Henry Zarrow School of Social Work

Spring 2015

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/14569



Hands On Testing (HOT) Pilot Project

Recommendations and Findings: Final Report

Spring 2015

Acknowledgements

The authors would like to thank the following individuals and organizations for their support of this project:

- The DHS Child Welfare Training staff, especially Guy Willis, Amanda Beard, Michele Roberts, Mark Carson and Carla Shackelford for all of the scheduling, coordination and logistical work needed to make the pilots run smoothly.
- The CWPEP payback program for financial support of this project.
- The SCOUT Simulation Center at OU Tulsa—particularly John Gaudet and Julie Roebeck for their partnership in developing and executing the simulated interviews

Table of Contents

Overview 3
What is the HOT test? 4
Purpose of project 4
Objective 1: Develop standardized measure/rubric for adult and child interview 5
Objective 2: Test reliability/validity of measure 6
Objective 3: piloting standardized actors and simulation for testing 20
Recommendations 22
Summary 27
References 28
Final Rubrics 29
Appendix A: HOT Testing Rubric Adult– Revised Final Version 29
Appendix B: HOT Testing Rubric Child – Revised Final Version 32

List of Tables & Figures

Table 1: Chronbach's Alpha on Child and Adult HOT rubrics: Final Version 10
Table 2: Description of new measure, categories, items, scale 11
Table 3: Ranked List Items from Adult Assessment 12
Table 4: Ranked Items Child Assessment 13
Table 5: Ratings of Agreement with the Following Questions 14
Table 6: Employment Area of Evaluators (old HOT data) 16
Table 7: Type of Child Welfare Interview (old HOT data) 16
Table 8: Communication Skills (old HOT data) 17
Table 9: Information Gathering Skills (old HOT data) 17
Table 10: Rapport Building Skills (old HOT data) 17
Table 11: Areas of Improvement, Coded into Categories 18
Table 12: Areas of Improvement, Subcategories & Example Quotes 19

Figure 1: CW Assessment of Validity of Current HOT test 3
Figure 2: Overview of the Project 4
Figure 3: Scale Components Adult Instrument 10
Figure 4: Scale Components Child Instrument 10
Figure 5: Distribution of Evaluators of Old HOT by Region 15

Overview

As stated in the Oklahoma Pinnacle Plan, it is critical to employ a competent, committed, trained and resourced Child Welfare workforce. ***Pinnacle Plan Point 3 Initiative 7 states: Effective September 1, 2012, training for new Child Welfare Specialists requires successful completion of a performance competency evaluation prior to caseload assignment.*** The Oklahoma Department of Human Services (DHS) Child Welfare Division is responsible for training new child welfare specialists.

DHS in partnership with The University of Oklahoma Health Sciences Center, and The University of Oklahoma Anne & Henry Zarrow School of Social Work (SSW) developed Hands on Testing (HOT) to provide the competency testing. DHS states that the purpose of this evaluation is to assess critical skills needed for child welfare workers to achieve positive outcomes for children and families. Hands on Testing is comprised of four skills-based components structured to support the successful execution of the Pinnacle Plan. Performance reached through Hands on Testing is directly linked to Child Welfare Services' goal of improving outcomes for the children and families we serve. Participants will achieve *successful completion* of Hands on Testing once they have demonstrated skills on all four components: a child interview, an adult interview, a safety assessment, and utilization of "KIDS" which is Oklahoma's statewide automated child welfare information system. HOT was implemented in Oklahoma in September, 2012.

The SSW was granted two "Payback Grants" totaling $28,000 from the Child Welfare Professional Enhancement Program (CWPEP) for the purpose of enhancing the simulated interview component of HOT and increasing the objectivity of the grading process. Payback Grants further the mission of CWPEP in enhancing and supporting Oklahoma's public child welfare workforce. Funding for Payback Grants comes from monies collected from former CWPEP students who do not honor their employment obligation to DHS and must repay the funds that were expended on their behalf. The federal title IV-E portion of the repayment fund is returned to DHS and the state share is placed in the CWPEP Payback Fund at OU. Payback Grants must be approved by the CWPEP participating universities, by OU Sponsored Programs and by DHS Child Welfare Services.

The two components of HOT that are the subjects of this project are the simulated interviews with a child and an adult. In the existing system, scenarios have been developed by the CW training staff and various staff members are used as actors. Graders are CW Supervisors who have been briefed on the scenarios and given a check list of what to evaluate during the interviews. The grader sits in the room with the student and actor, observes the interview and gives a pass (complete) or fail (incomplete) grade. Although response rates have been low, (rates during the first year were thirty percent (30%) and rates for the second year stand at twenty percent (20%), a significant number of CW Supervisors of workers completing HOT indicated that they do not believe that the system is a fair or valid way to evaluate the new workers because of the level of subjectivity in the grading process. Comments indicate that the level of subjectivity is not controlled because only one supervisor is responsible for observing and rating each interview and there is no well-defined grading rubric to guide the graders' decision making.

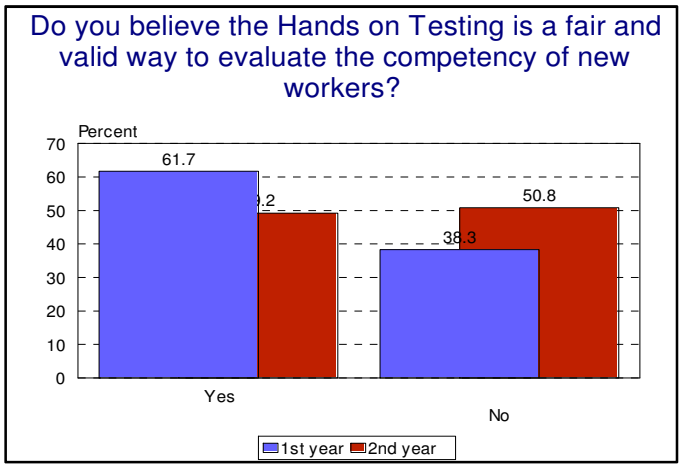


Figure 1: CW Assessment of Validity of Current HOT test

What is the HOT test?

The DHS "Hands On Testing Workers' Handbook" describes the interview components of HOT testing as follows:

Component One – Interview of a Child

During this portion of the competency evaluation, participants conduct an interview with an adult actor playing a child. The interview is based on the participant's primary role within Child Welfare: Child Protective Services, Permanency Planning, Foster Care, Adoptions, or Hotline. The interview is conducted one-on-one with an adult actor portraying a child in a neutral setting. Participants are expected to engage the child and gather information regarding safety, permanency, and well-being. Participants are provided a specific time frame of one hour to prepare for and conduct the interview.

Component Two – Interview with an Adult

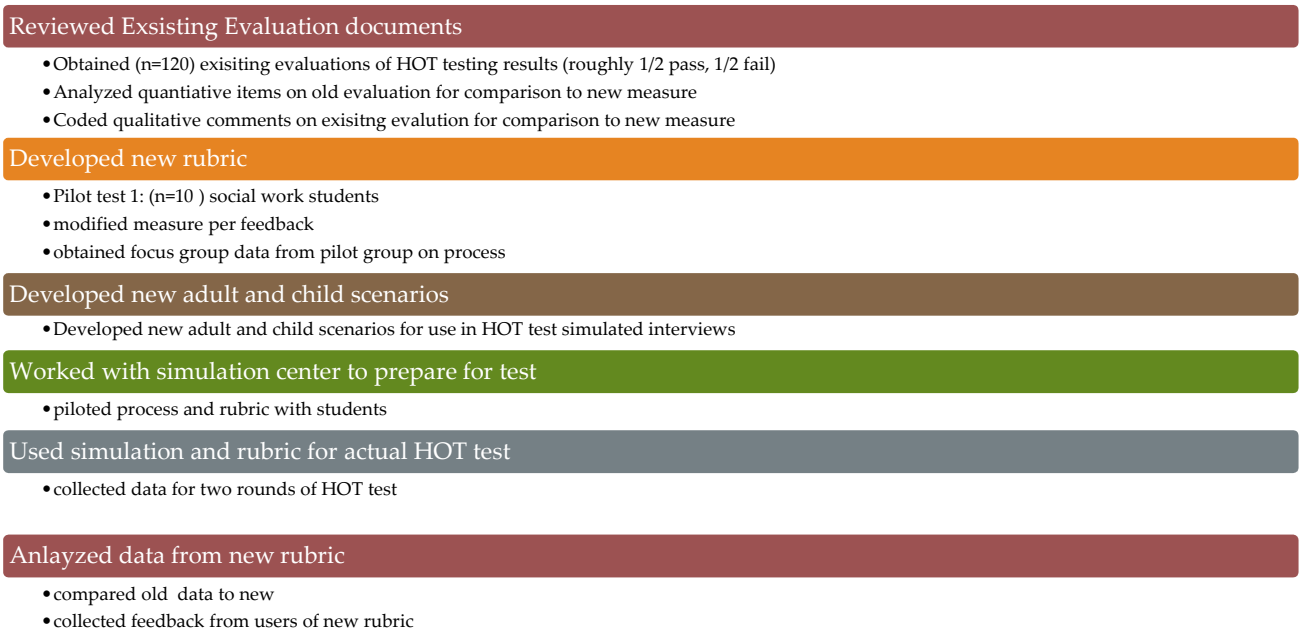
During this portion of the competency evaluation, participants conduct an interview with an adult actor. The interview is based on the participant's primary role within Child Welfare: Child Protective Services, Permanency Planning, Foster Care, Swift Adoptions, or Hotline. The interview is conducted one-on-one with an actor portraying the parent of a child involved in a Child Welfare case. The parent actor being interviewed is the parent of the child previously interviewed during component one of the competency evaluation. Participants are expected to engage the parent and gather information regarding safety, permanency, and well-being. Participants are provided a specific time frame of one hour to prepare for and conduct the interview.

Purpose of project

This project involved three main objectives and sub-goals as listed below:

Objective 1: Develop standardized measures for the HOT test (child and adult interview)
Objective 2: Test reliability/validity of the newly developed measures
Objective 3: Pilot-test using standardized actors and simulation for HOT testing

Figure 2: Overview of the Project



Objective 1: Develop standardized measure/rubric for adult and child interview

The purpose of this objective was to develop new tools (child and adult interviews) for use in the DHS Hands-on Testing for new child welfare employees. It was the aim of the project to create new tools that would 1) provide a more nuanced method of rating the skill level of testers in comparison to the existing checklist; and 2) have established reliability and validity based on pilot use of the tools, psychometric testing, and further revision for standardization. The process of developing the new measure for use in the HOT evaluation process with DHS employees followed the model for standardized tool creation (Barry, Chaney, Stellefson, and Chaney, 2011). It involved a decision-making process of systematic review of existing materials, establishing understanding of the phenomenon of interest, development of key survey constructs, a recursive process of item development and elicitation of feedback from experts, followed by piloting, testing for content validity and other psychometric properties of the tool, and evaluation of scale items using statistical methods such as factor analysis.

This portion of the project was completed by OU Assistant Professor, Dr. Kirsten Havig, whose expertise in the area provided an understanding of the phenomenon of interest (a client interview of either an adult or a child by a child welfare worker) without the need for additional research into the topic. Dr. Havig has been a social work educator since 2006, teaching clinical courses including Interaction Skills at the graduate level. She has also worked as a licensed therapist primarily with the child welfare population and has completed special training courses in forensic interviewing and other key areas of knowledge related to this project.

The initial step, review of existing materials, began first with the existing tools previously utilized in the hands-on testing of new child welfare workers, the Adult and Child Interview Checklists. The existing Adult and Child Interview Checklists allowed the reviewer to indicate whether or not a skill was demonstrated at a nominal level (yes/no) using a checkmark. Testers were required to achieve a benchmark number of checked items to pass (e.g. four of eight skills for “rapport” and “communication” and three of six skills for “assessing safety” indicated a passing score, in addition to additional items such as “introduced self, the purpose of the interview”, etc.). Next, a literature review was conducted using University of Oklahoma databases (Academic Search Premiere; Social Work Abstracts; PsychInfo) to search for related existing tools using key words such as, “measurement”, “interview skill”, “interaction skills”, “inventory”, “index”, “tool”, and others. No existing tools encompassing both the items from the Interview Checklist and items deemed critical to an interview of this nature were found. The literature search did reveal common elements found to be essential to interviewing skills including the importance of rapport-building, the benefits and drawbacks of open- vs. closed-ended questions, use of nonverbal cues, and crafting non-leading interactions with children (Carpetto, 2008; Carrillo, Gallant, & Thyer, 1993; Cheung, 2008; Friend, 2009; Pike, Bennett, & Chang, 2004; Stevenson, Leung, & Cheung, 1992). Next, an informal request was communicated to social work educators nationwide via a large listserv for examples of rubrics, indexes, or other tools that measure interviewing skills in a social work setting. Several examples were received and reviewed; none had been published, but were tools utilized primarily by social work educators at the baccalaureate and master’s levels to assess skill levels in classroom role play or simulation exercises (Zastrow, n.d.). Although no instrument was found that encompassed all the areas and skills deemed essential for HOT evaluation, this process did allow the team to ensure that the new tool being produced was exhaustive in comparison to others. Additionally, an internet search targeting sites including the Child Welfare Information Gateway and regional public child welfare departments revealed no additional helpful materials.

The first versions of the new HOT Rubric (Adult and Child) were then created based on this review of the literature and existing tools, as well as the key interaction and interviewing skills that are standard elements of social work coursework,

and the original tool. The three areas evaluated using the original HOT Interview Checklist were maintained for continuity in the new tools (rapport-building, communication skills, assessing safety). Items from the original Interview Checklists were rephrased or collapsed when multiple items seemed to overlap (e.g., respect and nonjudgmental stance combined). Additional items were added to the appropriate areas when a key skill was found to be absent (e.g., established clear expectations of next steps and follow up). Finally, a major task during rubric development was to clearly operationalize each item. Doing so not only helped to ensure exhaustive and mutually-exclusive items, but was a critical step in creating an instrument with a high level of reliability. By operationalizing each skill area, raters can be sure they are looking for and scoring based on the same specific, observable behaviors demonstrated by testers (e.g., "avoided leading and coercive questions" from the Interview Checklist became, "avoided leading and coercive questions – gives choices, does not suggest one response over another, etc.").

After engaging in the review of literature and existing instruments, as well as the recursive process described above for refining rubric items, the final version of both the Adult and Child HOT Rubrics were completed. Each rubric reflects the same core set of skills, organized into three primary areas: **Rapport-building** (eight items); **Communication and Information-gathering skills** (12 items); and **Safety Assessment and Ending Skills** (seven items). The Qualtrics electronic version of the tool allows for each subsection to be scored as it is completed so that skill level for each may be clearly identified and reported. A benchmark for a passing score was not identified and will be at the discretion of DHS for use with child welfare workers. Additionally, a summed total score is provided that includes all three skill areas. As noted above, there are also text boxes for raters to include comments that will help to inform the rated scores provided and to provide specific feedback about strengths, concerns, and suggestions for improvement.

Scoring on the Adult and Child HOT Rubrics was constructed using a five-item Likert-type scale from 0-4. The scores are defined as follows: 0 = Poor, no evidence of skill, not competent; 1 = Fair, lacks clear evidence of skill, limited understanding; 2 = Good, some evidence of skill, emerging competence; 3 = Very good, clear evidence of skill, competent; and 4 = Excellent, ample evidence of skill, very competent.

Objective 2: Test reliability/validity of measure

After the rubrics were constructed, they were tested in two phases. Phase 1 included a pilot test of the simulation logistics and new rubrics with social work students. Adjustments were made to the rubric based on findings from the pilot data and feedback from raters. Phase 2 included using new rubrics with actual child welfare workers during an actual HOT test. . In order to compare the old evaluation measures with the new rubric, evaluation data from the existing HOT testing measure was obtained and analyzed. This analysis allowed for a comparison between existing and new measures of the HOT test. Each of these is described below.

Phase 1: Pilot

Ten CWPEP practicum students, both BSW and MSW from the Norman and Tulsa campuses, engaged in interviews using simulation and new rubrics. Each interview was videotaped, additional copies were made and were distributed to a group of trained graders who were not CW Supervisors, but were MSW professionals with extensive experience in child welfare services. The graders utilized a draft rubric to review the interviews and the ratings were entered into a Qualtics survey for analysis and the rubric was revised based on reviewer input.

Once the pilot version of both the Child and Adult HOT rubrics were created, all items along with the new scoring system (described below), were entered into an electronic survey by the raters. The first round of piloting the new tools also involved use of simulated clients for recorded interview tests who had received training on their client scenarios as well as

the key skills that the DHS workers should demonstrate. Each rater watched 10-12 interviews and scored the testers using the new Adult and Child HOT Rubrics. Additionally, a survey was created on Qualtrics that was linked to the rubrics for raters to provide feedback about the process and instrument used in the pilot round. The research team met to discuss the feedback and experiences using the new rubric; at that time several items were removed or collapsed and initial factor analysis was done. Additionally, items were added to both the Adult and Child HOT Rubrics that allow for raters to include more qualitative information and summative conclusions about the tester (e.g., worker's overall readiness for the field; space to write worker strengths and any concerns). This will allow for individualized feedback, focus on worker strengths, and the recording of red flags in addition to the standardized, quantitative items on the tool. Feedback from the raters also pointed to the need to categorize testers and accompanying rubric depending on whether the context was permanency planning, etc. As a result, items were added to align with those specific areas and that allowed for evaluation of specific information-gathering.

The research team met to discuss the feedback and experiences using the new rubric, after an analysis of the data produced with it was conducted. In that initial statistical analysis, measuring the effectiveness of the rubric, several promising findings emerged. Among them were remarkably high Chronbach's Alpha scores for both the Adult (.92) and Child (.86) instruments. Chronbach's Alpha is a correlation coefficient measurement that assesses the interrater reliability of a data collection instrument, and the highest possible score is a 1. It is commonly accepted that any score above .70 suggested the instrument is reliable and measuring a phenomenon consistently among raters (George & Mallery, 2003). Additionally an investigation of the correlation of rubric items was used to assess the scale item independence, and a factor analysis was completed which showed the items on the scale did not cluster as individual sub-groups, but rather as a whole.

After this initial evaluation of the psychometrics of the rubric, several items were removed, collapsed, and clarified. Additionally and subsequent to rater feedback, items were added to both the Adult and Child HOT Rubrics that allow for raters to include more qualitative information and summative conclusions about the tester (e.g., worker's overall readiness for the field; space to write worker strengths and any concerns). This allowed for the rubrics to catch data based on individualized feedback, which could focus on worker strengths and the recording of red flags in addition to the standardized, quantitative items on the tool. Feedback from the raters also pointed to the need to categorize testers and accompanying rubric depending on whether the context was permanency planning, etc. As a result, items were added to align with those specific areas and that allowed for evaluation of specific information-gathering.

Feedback from Simulated Clients (actors)

Post student pilot, a debriefing was conducted with the actors who portrayed the simulated clients to gather additional data pertaining to their perceptions of workers' skills in rapport-building, communication, and information gathering. This discussion revealed key elements of the HOT simulation from the perspective of the "client" that reflect the critical skills assessed

Attending to person over task → *We know that you guys have a checklist and that there's certain information that you guys just, you have to have, cause if you don't have this information how can you help this person if you don't get it? And a lot of times, that, that becomes, you kind of get blinders on, (puts hands towards side of face) and you're like "I, I gotta get this information. I understand this kid's been through a lot, but, I, to help this kid, I've gotta get this information." And with those blinders on, you may forget about your body language, about what you're doing, your tone of voice. You may just get into that zone of, "I need to gather this information."…It makes me feel like you're not even making a connection at all. It just makes me feel like I'm just another number. Another one in a long list of kids you gotta see today and get all this information. So, body language and the information-gathering mentality - those are the two things that I took away from this…*

by the rubrics. The simulated clients highlighted the importance of non-verbal communication; perceptions of workers' ability to demonstrate care for the person over the need to complete a particular task; the importance of authentic interest in the client and empathy to the quality of the interaction; and the value of open, transparent interactions to the client's ability to trust. For example, one of the actors discussed the dilemma of **attending to the person over the task**.

From the perspective of the actor, it was apparent that the workers had 'checklists' and particularly information to gather. However, they also noted that a balance between information gathering and establishing rapport is an important skill. Similarly, another actor noted the importance of focusing on interest in the person over the 'investigation' commenting that "… *And so I really liked that because then my character was easier to kind of portray, "Okay, I'm a little more open, I'm a little less guarded." Because it wasn't investigating "What's going on with you? What's wrong with you?" It was, "So, what do you like to do? What is, what is your…".* This feedback speaks to the importance of treating the person being interviewed with respect and dignity. The actors were able to validate the need to address those being interviewed as complex and individualized people with valuable stories. This early data from the actors confirms later findings about the important balance between information gathering and establishing rapport throughout client contact.

It is still my life → *I had an interaction where I kept hearing um "my case this" or "your case this" over and over and over again, and that felt like a disconnect between me and that person because I felt like, "Yes, my life has now become a government case now" and I understand that. But you know, it is still my life and I know it's being condensed all into a file, but I as a character would appreciate if it wasn't referred to as "my case this" or "my case that" or at the very least, not repeated so many times that I'm constantly being told, "Oh, I'm just another file in a drawer somewhere with 1,500 other kids that have had similar experiences." Just refer to me as a person, even if I may be 6 or 16, age is not a factor here. You're a human being, and that's how I shall treat you regardless of your age.*

Phase 2: Testing rubrics using CORE participants

The initial plan was that Child Welfare Supervisors would be trained on the use of the rubric. To address the issue of subjectivity of the grader, it was planned that the interviews would be videotaped and mailed to three supervisors for grading and pass/fail determinations would be based on the average of the three grades. This would also allow the supervisor/graders to remain in their offices, which would save significant time and travel funds and perhaps would reduce the test anxiety of the HOT participant as there would be no grader in the room during the interview. However, it was determined that this would not be possible as it was essential to DHS that the HOT results be made available to the participants on the same day as they completed the testing. In addition, the graders needed to be available at the training site to debrief with the participants, particularly those who did not pass (ie, received an incomplete) one or both of the interviews.

The solution was to use MSW trained raters (many who rated in the pilot) to review and rate the performance of the CORE participants using the revised rubric. The CORE participants completed the interviews on four dates: May 12-13, 2014 and June 24-25, 2014. The interviews were videotaped for later review. However, the CW supervisors were present during the interview and utilized the previously used DHS checklist for assessment.

Reviewers

A group of content matter experts were selected to review the taped interviews and rate them utilizing the rubrics. The graders were MSW or above practitioners and professors who had extensive child welfare related experience, including previous experience as child welfare workers, supervisors and managers as well as providing consultation to child welfare supervisors (refer to list of reviewers below). A brief training was provided to the graders, as the rubrics were considered to be self-explanatory. Raters were all sent DVDs of the interviews they were to rate and all raters entered rubric scores into the electronic survey. Graders quickly become comfortable with the rubrics as they used them. A total of 55 CORE participant interviews were assessed by twelve raters.

Reviewers

Carol Clabo, MSW – retired DHS Child Welfare
Mary Grissom, MSW – SSW CWPEP, former DHS Child Welfare
Serena Hanson, MSW – SSW Community Faculty, former CASA and TFC
Kirsten Havig, MSW, PhD – SSW assistant professor
Allison Holeman, MSW – DHS Child Welfare
Cindy Howard, MSW – SSW Field Instructor and Liaison, former DHS Child Welfare
Aleece Mann, MSW – DHS Child Welfare
David McLeod, MSW, PhD - SSW assistant professor
Kathy Simms, MSW, - SSW Field Liaison & Clinical Specialist, former DHS Child Welfare
Linda Smith, MSW – SSW CWPEP, former DHS Child Welfare
Jon Trzcinski, MSW – CWPEP Clinical Specialist, former DHS Child Welfare
Elaine Waters, MSW SSW CWPEP

Results

After engaging in the review of literature and existing instruments, as well as the recursive process described above for refining rubric items, the final version of both the Adult and Child HOT Rubrics were completed. Each rubric reflects the same core set of skills, organized into three primary areas: Rapport-building (eight items); Communication and Information-gathering skills (twelve items); and Safety Assessment and Ending Skills (seven items). The Qualtrics electronic version of the tool allows for each subsection to be scored as it is completed so that skill level for each may be clearly identified and reported. A benchmark for a passing score was not identified and will be at the discretion of DHS for use with child welfare workers. Additionally, a summed total score is provided that includes all three skill areas. As noted above, there are also text boxes for raters to include comments that will help to inform the rated scores provided and to provide specific feedback about strengths, concerns, and suggestions for improvement.

Scoring on the Adult and Child HOT Rubrics was constructed using a five-item Likert-type scale from 0-4. The scores are defined as follows: 0 = Poor, no evidence of skill, not competent; 1 = Fair, lacks clear evidence of skill, limited understanding; 2 = Good, some evidence of skill, emerging competence; 3 = Very good, clear evidence of skill, competent; and 4 = Excellent, ample evidence of skill, very competent.

Tables 3 and 4 detail ranked lists of the average scores for each of the items from the second rubric pilot. Appendix 1 shows the scores from the adult rubric and Appendix 2 details the scores from the child rubric. Both are based on (n=125) rater observations and the findings suggest the HOT trainees consistently showed the highest levels of proficiency in *Rapport Building*, with the next highest levels of proficiency being observed in *Communication and Information Gathering Skills*, and the least proficiency being observed in *Safety Assessment and Ending Skills.*

Twelve raters used the electronic version of the rubric, on the Qualtrics site, to evaluate a total of 55 participants. Additional survey data and feedback was again elicited from the raters and discussed in a research team meeting. After this second pilot the instrument's psychometric properties showed to be even stronger than the first version. The adult version of the instrument showed a Chronbach's Alpha score of .977 and the child version produced a score of .978. These demonstrate a very high correlation of variance between raters and suggest that the instrument is so strong in fact, that the number of items on it could likely be reduced and it would still produce significant interrater reliability. As is seen in the table below there were no observed weaknesses when alpha scores were assessed by individual section. Even in

the final section (Overall Ratings) where the assessments could be seen as somewhat more subjective, the raters showed a tendency to follow similar patterns in evaluation.

Table 1: Chronbach's Alpha on Child and Adult HOT rubrics: Final Version

	Child α	Adult α
Full Instrument	.978 (.001)	.977 (.001)
Rapport Building Skills	.933 (.001)	.928 (.001)
Communication and Information Gathering Skills	.952 (.001)	.950 (.001)
Safety Assessment and Ending Skills	.924 (.001)	.920 (.001)
Overall Ratings	.913 (.022)	.925 (.039)

Note: significance level of p< is included in parentheses next to the alpha score for each category

In short, the final versions of the adult and child rubrics proved to be remarkably strong in terms of interrater reliability and internal consistency. These findings were consistent across all sections of the instrument. A varimax rotated principle components factor analysis was also conducted to assess if the separate sections of the instrument would load as freestanding and autonomous components of the scale. The visual representation of that procedure, as seen in the scree plots of Figures 1 & 2, shows that from a mathematical analysis of variance the items on the full instrument load as one complete scale of overall worker performance in the interview rather than three independent components assessing rapport building, communication, and safety assessment.

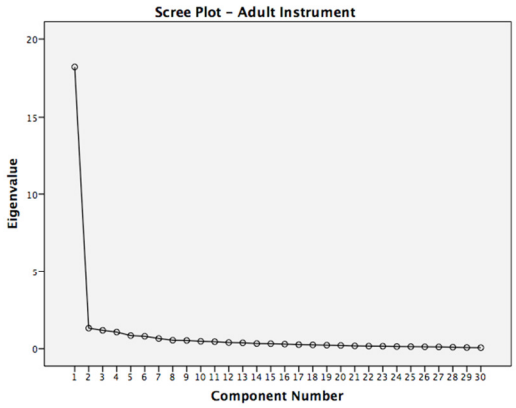


Figure 3: Scale Components Adult Instrument

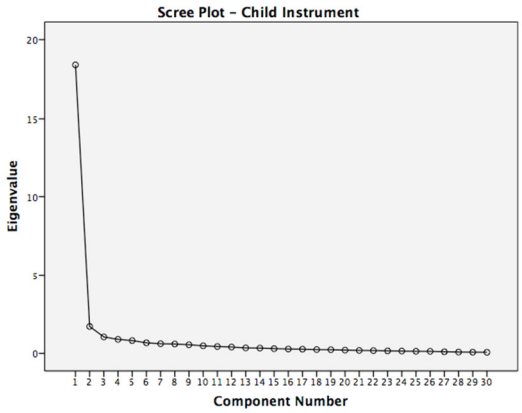


Figure 4: Scale Components Child Instrument

The highest scoring items from both the adult and child interviews still only showed mean scores of 2.56 and 2.67 respectively. Based on the scoring criteria (0-4) this suggests that HOT trainees were, at best, demonstrating emerging competence in some areas. In the lowest overall area of *Safety Assessment and Ending Skills* the average mean score on the child interviews was 2.07 and an even lower score of 1.97 was observed from the interactions with adults. These scores suggest that raters observed that on average HOT trainees appeared to be showing some evidence of emerging skill, but that they were dangerously close to presenting with a lack of skill or limited understanding of safety assessment.

Inter-item correlations were used to assess for possibilities related to item reduction in both the child and adult instruments. As the high alpha scores for internal consistency may suggest, there are multiple inter-item correlations that show how items could be reduced if scale length were a concern to program developers. Table 2 (following page) details which items could be considered for merger and sorts the findings by child and adult instruments. As is seen, most specifically with the adult instrument, the domain areas (rapport, communication, and safety) may be measuring a single

phenomenon as much as the independent constructs related by the individual items. While these items do help to give specificity to the exact behaviors observed during HOT testing, the high inter-item correlations suggest that they are in fact measuring the same or at least similar constructs across items. For example, in the communication section of the adult instrument significant correlations are observed between items related to specific and open-ended questioning, probing and exploration in questioning, avoiding coercive questioning, and others. It should be assessed to determine if these specific items are needed for the HOT assessment, or if a more broadly defined item such as "employs quality interview techniques" could serve as a single inclusive item that could take the place of all three and work to shorten the overall length of the instrument while still assessing for worker competence. Table 2 lists several of the more pronounced item correlations that could be taken under consideration if shorter instrumentation were desired.

Table 2: Description of new measure, categories, items, scale

Item Correlations on Child and Adult Instruments		
	Child Instrument	**Adult Instrument**
Rapport	• Items 1 (Introductions) & 2 (Established clear expectations) correlate at .76 Item 3 (Respectful and non-judgmental) is also correlated with item 1 at .716 and the three items could be merged into a single one.	• Items 4 (Demonstrated empathy) and 8 (Use of appropriate language) correlated at .758 and could be merged into a single item. • Items 3 (Respectful and non-judgmental) and 5 (Non-threatening) correlated at .804 and could be merged.
Communication	• Items 1 (Used focused, open ended questions) and 2 (Use of silence) were correlated at .754. • Item 1 also correlated with item 6 (using language mirroring and client-defined terms) at .706 and item 9 (Avoided leading and coercive questions) at .783. • Items 7 (Able to redirect and/or keep interview focused) and 8 (Employed probing and exploration questions) correlated at .848.	• Item 1 (Used focused, open ended questions) correlated with item 8 (Employed probing and exploration questions) at .721 and with item 9 (Avoided leading and coercive questions) at .727. • Item 2 (Use of silence) correlated with item 9 as well at .773. • Item 1 also correlated with item 12 (Checked to assess for understanding) at .715. • Item 2 (above) was also correlated with item 11 (Demonstrated ability to question inconsistencies) at .791. • Item 6 (using language mirroring and client-defined terms) correlated with item 4 (Appropriate use of paraphrase or summarization) at .737 and item 5 (Use of developmentally-appropriate terms) at .711. • Item 6 also correlated with item 9 (Avoided leading and coercive questions) at .748. • Item 12 (Checked in with client to assess for understanding) also correlated with Item 3 (Active listening at .727 and item 4 (Appropriate use of paraphrase) at .739.
Safety	• Items 1 (Gathered necessary information) and 2 (Explored safety issue from client perspective) correlated at .789. Item 2 was also highly correlated with item 5 (Used probing questions) at .749. • Item 3 (Established clear expectations of next steps) showed a very high correlation with item 7 (Follow up plan/next steps made clear) at .904.	• Item 1 (Gathered necessary information) correlated with item 2 (Explored safety issue from client perspective) at .758 and with item 5 (Used probing questions) at .816. Item 5 also correlated with item 2 (Explored safety issue from client perspective) at .804. Items 1 and 2 had low correlation. • Item 3 (Established clear expectations of next steps) and item 4 (Gauged PRFC's understanding of safely concern) correlated at .724. Item 3 also correlated with item 7 (Follow up plan/next steps made clear) at .724.

Tables 3 and 4 (below) list, in rank order, the mean scores from rubric items. The scores are listed in descending order, to give indicate which items participants tended to score highest and lower on. Thematically it can be seen in the tables how on both adult and child instruments, participants tended to rank higher in competence as related to rapport building, followed by communication gathering skills. On both adult and child cases the data suggests participants ranked lowest in

safety assessment. It should also be noted how on a scale of 0-4 the highest scores on both assessments were in the mid 2's. This suggests that at best reviewers are ranking participants with emerging competence on average, rather than as exhibiting clear or ample evidence of skill.

Table 3: Ranked List Items from Adult Assessment

	Mean	Std. Deviation	N
Rapport Building Skills: Non-threatening, approachable demeanor (receptive, open posture, eye contact, warmth/welcoming voice tone, etc.)	2.56	0.766	125
Rapport Building Skills: Respectful, nonjudgmental stance (refrained from blaming; active listening, ect.)	2.56	0.787	125
Rapport Building Skills: Professionalism (courteous; appropriate presentation; use of professional language; boundaries; ect.)	2.48	0.801	124
Communication and Information Gathering Skills: Active listening - eye contact, neutral responses (head nod, attentive posture, etc.)	2.38	0.904	125
Rapport Building Skills: Ability to use age-appropriate language and to adjust interaction to client (due to age, mental or emotional state, etc.)	2.33	0.814	124
Rapport Building Skills: Demonstrated empathy (able to accurately reflect feelings of client; indicated understanding and interest; expressed concern; acknowledged client)	2.27	0.937	125
Communication and Information Gathering Skills: Able to redirect client and/or keep interview focused on its stated purpose	2.27	0.807	125
Rapport Building Skills: Introductions- clearly identified self and role; addressed client appropriately; used names	2.23	0.903	124
Communication and Information Gathering Skills: Used focused, open ended questions (vs. yes/no or asking more than one question at a time)	2.22	0.771	125
Communication and Information Gathering Skills: Employed probing and exploration questions to gather information	2.21	0.845	125
Communication and Information Gathering Skills: Allowed client to ask questions/provided appropriate responses to question	2.18	0.807	125
Communication and Information Gathering Skills: Avoided leading and coercive questions (gives choices, does not suggest one response over another, etc.)	2.18	0.807	125
Communication and Information Gathering Skills: Use of developmentally appropriate terms; refrain from use of jargon; explained unfamiliar terms	2.16	0.723	125
Safety Assessment and Ending Skills: Refrained from making promises or providing premature predictions about the future	2.13	0.967	125
Communication and Information Gathering Skills: Use of silence - allowed client to speak, gather thoughts, did not interrupt	2.08	0.782	124
Safety Assessment and Ending Skills: Used probing questions to gather further information about safety (context and nature of concern)	2.08	0.829	125
Safety Assessment and Ending Skills: Validated client experiences and concerns - received information with care and concern	2.06	0.957	125
Communication and Information Gathering Skills: Appropriate use of paraphrase or summarization to reflect content of client statements	2.02	0.801	124
Rapport Building Skills: Established clear expectations; explained purpose and process of interview; explained confidentiality	1.99	0.897	124

Communication and Information Gathering Skills: Ability to join with client by using language mirroring and client defined terms	1.99	0.781	124
Rapport Building Skills: Strengths-based attitude (highlighted client abilities and strengths in addition to concerns)	1.97	0.892	124
Safety Assessment and Ending Skills: Gathered necessary information for assessing child safety	1.97	0.806	124
Safety Assessment and Ending Skills: Follow up plan/next steps made clear	1.93	0.93	124
Safety Assessment and Ending Skills: Explored safety issue from client perspective	1.86	0.939	125
Communication and Information Gathering Skills: Checked in with client during interview to assess for understanding	1.8	0.803	125
Communication and Information Gathering Skills: Demonstrated ability to question inconsistencies or confront in respectful manner	1.8	0.843	114
Safety Assessment and Ending Skills: Gauged Clients understanding of safety concern(s) before terminating interview - for example, asked for client summary of critical information; rephrased if lack of client understanding is evident	1.78	0.912	125

**scale: 0 = Poor, no evidence of skill, not competent; 1 = Fair, lacks clear evidence of skill, limited understanding; 2 = Good, some evidence of skill, emerging competence; 3 = Very good, clear evidence of skill, competent; and 4 = Excellent, ample evidence of skill, very competent.*

Table 4: Ranked Items Child Assessment

	Mean	Std. Deviation	N
Rapport Building Skills: Non-threatening, approachable demeanor (receptive, open posture, eye contact, warmth/welcoming voice tone, etc.)	2.67	0.765	123
Rapport Building Skills: Professionalism (courteous; appropriate presentation; use of professional language; boundaries; ect.)	2.6	0.744	123
Rapport Building Skills: Respectful, nonjudgmental stance (refrained from blaming; active listening, ect.)	2.58	0.665	123
Communication and Information Gathering Skills: Active listening - eye contact, neutral responses (head nod, attentive posture, etc.)	2.52	0.833	123
Rapport Building Skills: Demonstrated empathy (able to accurately reflect feelings of client; indicated understanding and interest; expressed concern; acknowledged client)	2.41	0.896	123
Communication and Information Gathering Skills: Employed probing and exploration questions to gather information	2.38	0.784	123
Safety Assessment and Ending Skills: Validated client experiences and concerns - received information with care and concern	2.3	0.905	123
Rapport Building Skills: Ability to use age-appropriate language and to adjust interaction to client (due to age, mental or emotional state, etc.)	2.29	0.875	123
Communication and Information Gathering Skills: Use of developmentally appropriate terms; refrain from use of jargon; explained unfamiliar terms	2.28	0.782	123
Communication and Information Gathering Skills: Able to redirect dinette and/or keep interview focused on its stated purpose	2.26	0.663	123
Communication and Information Gathering Skills: Allowed client to ask questions/provided appropriate responses to question	2.26	0.871	121
Communication and Information Gathering Skills: Used focused, open ended questions (vs. yes/no or asking more than one question at a time)	2.25	0.775	122
Communication and Information Gathering Skills: Avoided leading and coercive questions (gives choices, does not suggest one response over another, etc.)	2.22	0.815	123
Rapport Building Skills: Introductions- clearly identified self and role; addressed client appropriately; used names	2.19	0.85	117

Communication and Information Gathering Skills: Appropriate use of paraphrase or summarization to reflect content of client statements	2.19	0.813	123
Rapport Building Skills: Strengths-based attitude (highlighted client abilities and strengths in addition to concerns)	2.15	0.893	124
Communication and Information Gathering Skills: Ability to join with client by using language mirroring and client defined terms	2.14	0.785	122
Safety Assessment and Ending Skills: Explored safety issue from client perspective	2.11	0.977	122
Safety Assessment and Ending Skills: Used probing questions to gather further information about safety (context and nature of concern)	2.1	0.882	123
Safety Assessment and Ending Skills: Gathered necessary information for assessing child safety	2.06	0.852	123
Safety Assessment and Ending Skills: Refrained from making promises or providing premature predictions about the future	2.04	0.876	122
Communication and Information Gathering Skills: Use of silence - allowed client to speak, gather thoughts, did not interrupt	2.01	0.815	123
Communication and Information Gathering Skills: Checked in with client during interview to assess for understanding	1.96	0.743	122
Safety Assessment and Ending Skills: Follow up plan/next steps made clear	1.96	0.931	122
Safety Assessment and Ending Skills: Gauged child's understanding of safety concern(s) before terminating interview - for example, asked for client to summary of critical information; rephrased if lack of client understanding is evident	1.95	0.978	122
Rapport Building Skills: Established clear expectations; explained purpose and process of interview; explained confidentiality	1.91	0.837	118
Communication and Information Gathering Skills: Demonstrated ability to question inconsistencies or confront in respectful manner	1.87	0.81	111

Raters Assessment of the new HOT Evaluation Rubric

A survey was sent to the evaluators to measure their perceptions of using the new evaluation rubric to assess worker performance during the HOT testing. Overall perceptions were generally favorable. In the table below a likert scale of 1-5 (ranging from strongly disagree to strongly agree) was used to assess rater agreement with the following statements:

Table 5: Ratings of Agreement with the Following Questions

Question	Mean (SD)
The HOT Evaluate Rubric included the essential skills related to **rapport building** that are required for and effective interview with an **adult** client.	4 (0)
The HOT Evaluate Rubric included the essential skills related to **rapport building** that are required for and effective interview with a **child**.	3.8 (.63)
The HOT Evaluate Rubric included the essential skills related to **communication and Information gathering skills** that are required for and effective interview with an **adult** client.	3.8 (.42)
The HOT Evaluate Rubric included the essential skills related to **communication and Information gathering skills** that are required for and effective interview with a **child**.	3.7 (.48)
The HOT Evaluate Rubric included the essential skills related to **safety assessment and ending skills** that are required for and effective interview with an **adult** client.	3.3 (1.06)
The HOT Evaluate Rubric included the essential skills related to **safety assessment and ending skills** that are required for and effective interview with an **adult** client.	3.3 (1.06)

The process of scoring the HOT Evaluate Rubric was clear.	3.09 (1.14)
I feel that the HOT Evaluate Rubric will provide an accurate score reflecting participant's actual interviewing skillset.	2.64 (1.03)

As is reflected in the table above, one general theme from throughout the evaluation continues. Evaluators showed far more consensus in their perceptions of the ability to capture data related to rapport building and communication, and were far more diverse in their perceptions on the ability of the form to adequately capture safety assessment. Additionally evaluators showed ambivalence about the clarity of the rubric, and were not convinced that it could be a consistent evaluative tool in measuring the participants "interviewing skillset."

Evaluators also provided qualitative feedback on the rubric. When asked what additional skills they believed were not covered by the rubric, the evaluators noted that some questions felt repetitive, and that safety questions were too broad. Other concerns included that there were not enough ways to count against someone for not assessing safety and that areas were needed to insert qualitative feedback. When evaluators were asked to add any additional comments they listed the following concerns:

- Include a mechanism to document when a specific behavior was not observed.
- Provide examples for items based on the categories. For example, provide a range of worker behaviors that would be rated as competent versus emerging competence or poor skill.
- Overall score does not provide level of specificity needed for feedback.
- Separate out evaluation of safety versus basic interviewing skills.

Comparison of current HOT assessment with new rubric

Data from previous HOT testing was obtained from DHS for comparison with findings from the newly constructed rubric. This was a challenging endeavor, as the HOT data from previous tests was different from year to year and from type of interview. However, we obtained approximately (n=120) old tests, and organized the qualitative comments, yes/no comments from previous tests for comparison. Results are described below.

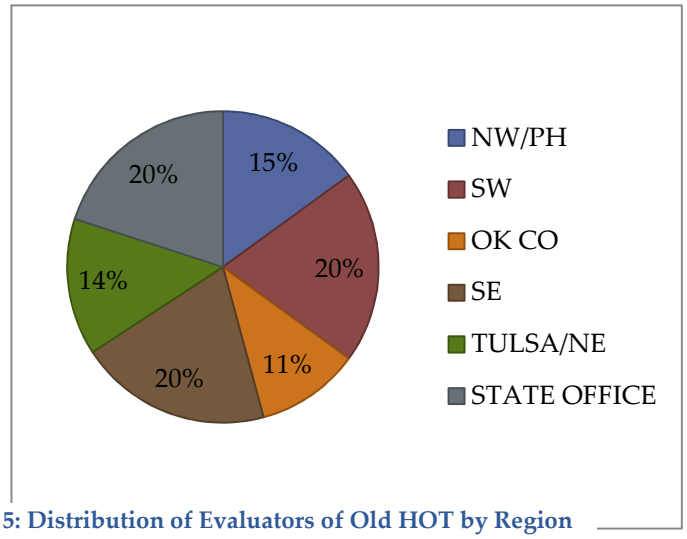


Figure 5: Distribution of Evaluators of Old HOT by Region

Data on the Evaluators

The evaluators of the previous HOT tests were located all over the state in a variety of child welfare settings. Thirty-two percent of the evaluators indicated that they had a Master's degree (19% with a MSW specifically) and sixty seven percent indicated a BA degree (7.5% with a BSW). On average the evaluators were quite experienced reporting an average of 11.3 years of work in child welfare and 6 years as a supervisor specifically.

Table 6: Employment Area of Evaluators (old HOT data)

	N	%
CPS Child Protective Services	25	20.8
PP Permanency Planning	32	26.7
COMP Comprehensive	23	19.2
DIST MGT District Management	1	.8
OCS Liaison	1	.8
CWS Child Welfare Supervisors	1	.8
CWS-DD Child Welfare District Director	1	.8
CWS Child Welfare Services	8	6.7
FC Foster Care	7	5.8
ADOPT Adoptions	8	6.7
CWS – PFR Child Welfare Services Programs Field Representative	10	8.3
CWS - Program Staff	2	1.7
CWS - FMG	1	.8
Total	120	100.0

As previously noted, 120 old HOT tests were collected and approximately half were successful HOT tests and the other half were failures. Seventy-five percent of the interviews were with adults (n=90) and twenty-five percent (n=30) interviews were child interviews. Table 6 indicates the type of child welfare interview.

Table 7: Type of Child Welfare Interview (old HOT data)

	n	%
Adult CPS	17	14.2
Adult Perm Plan/Adult PP	13	10.8
Adult Interview (not specific)	30	25.0
Child CPS	12	10.0
Child Permanency Planning	12	10.0
Child Foster Care	1	.8
Child Interview (not specific)	35	29.2
Total	120	100.0

As noted previously, there were multiple forms previously used to assess the past CORE participants so the data had to be organized for comparison to the new rubric. This was accomplished by conducting counts of yes/no items on the previous tests and coding qualitative comments into themes.

Tables (8 through 10) indicate the percentage of test takers who demonstrated a particular skill (YES = skill demonstrated; NO = skill not demonstrated) within three different categories: communication skills, information gathering skills and rapport building skills. Results are presented in rank order within the table, with the most highly demonstrated skills listed first. In these tables, rapport building appears to be the strongest area for participant workers.

This is followed by information gathering, which included some discussion of safety assessment. The area of lowest competence appears to be communication skills. This is a bit different from the outcomes of the newer assessment above and suggests that the two methods of data collection on worker performance may have been measuring different skills.

Table 8: Communication Skills (old HOT data)

COMMUNICATION SKILLS	N	% NO	% YES	
Doesn't overreact to anything the child says	23	13.0	87.0	**0-25% did not demonstrate**
Avoids making promises to the child	23	30.4	69.6	**26-49% did not demonstrate**
Tell the truth, refer to real events	22	31.8	68.2	
Clarifies child's responses when needed	23	34.8	65.2	
If DK, just say so, no guess	22	40.9	59.1	
You don't have to answer if you don't want to but I may ask again later	22	45.5	54.5	
Worker mentions… It's okay to correct me	22	50	50	**50% & above did not demonstrate**
Tell me if you don't understand the question	22	50	50	
Summarizes interview content	31	54.8	45.2	
Avoids leading or coercive questions	22	59.1	40.9	

Table 9: Information Gathering Skills (old HOT data)

Information gathering	N	% NO	% YES	
Asks if the child has questions or additional comments	23	13.0	87.0	**0-25% did not demonstrate**
Interview guide, contact guide	23	26.1	73.9	**26-49% did not demonstrate**
Gathers information relative to permanency & well-being	21	28.6	71.4	
Gathers information related to the area of concern	38	34.2	65.8	
Observes the physical and emotional condition of the child	23	34.8	65.2	
Uses open ended questions or tell me Invitations	23	43.5	56.5	
Gathers family demographic information	22	45.5	54.5	
Discusses the child's safety net	22	59.1	40.9	**50% & above did not demonstrate**

Table 10: Rapport Building Skills (old HOT data)

Rapport Building Skills	N	% NO	% YES	
Maintains professional boundaries	48	10.4	89.6	**0-25% did not demonstrate**
ID client, verify ID, greets appropriately	48	10.4	89.6	
Appears organized and prepared	23	13.0	87.0	
Introduces self and others and explains reason for interview	48	14.6	85.4	
Sets a positive tone for the interview	47	19.1	80.9	
Discusses role with any interview observers	19	21.1	78.9	

Questioning is child friendly and child focused	23	21.7	78.3	
assurance not in trouble	22	22.7	77.3	
Considers physical setting and arranges the room	21	23.8	76.2	
Addresses fears or concerns	40	25.0	75.0	
Attentive to own body language and demeanor	23	26.1	73.9	**26-49% did not demonstrate**
Thanks child for participating	23	26.1	73.9	
Appropriate rapport building	48	27.1	72.9	
Engages the child on a developmentally appropriate level	23	30.4	69.6	
Free or Directed Drawing, other tools	22	31.8	68.2	
Explains purpose of interview	31	32.3	67.7	
Gives appropriate explanations	23	39.1	60.9	
Explains next steps	39	48.7	51.3	

The old HOT test forms also included a place for evaluators to write comments related to three specific areas: areas for improvement, documentation, and general comments. Comments in each area were coded for general themes and then simple counts were used to indicate the most common themes within these four areas. The first set of comments analyzed were areas of improvement.

Table 11: Areas of Improvement, Coded into Categories

TOPIC	# of times this category mentioned in evaluator comments
1. Assess risk	56
2. Follow up specific questions	49
3. Explain purpose of interview	44
4. Gather information	38
5. Pace of interview	33
6. Use of open ended questions	32
7. Engagement	31
8. Rapport building	27
9. Use of tools	15
10. Disclosure	14
11. Prepared for interview	13
12. Provide or explain resources	11
13. Avoid bias in interview	5

Clearly the assessment of risk (over ½) was the most common area noted for improvement, followed by follow-up questions, explaining purpose of interview and gathering information. The following table provides an example quote from each major category of areas noted for improvement.

Table 12: Areas of Improvement, Subcategories & Example Quotes

TOPIC	subcategories	Example
Assess risk	56	*Could have talked a little more about mom's functioning while on so many Loritabs and how that may affect caring for her son (& if co sleeping could be dangerous while on meds)*
Follow up specific questions	Failed to follow up with question (37), failed to clarify or get clarification (12)	*Ask child to define "safe." Need to explain why taking notes. Used word "normally" - remember this is subjective and a child may not be able to define. Asked one leading question regarding medication.*
Explain purpose of interview	Did not introduce self (6), did not explain reason for visit (15), did not explain process (24), multiple mistakes (1)	*Always explain why/what you're writing, it builds rapport & puts people at ease.*
Gather information	38	*Didn't ask about pill bottles on the floor. That was a concern in the referral, but the form also says the worker did not observe safety hazards.*
Pace of interview	Too slow or fast (20), misguided direction 10)	Asked too many questions in a row – let them answer. Interview was very formal & direct. Not a lot of open-ended questions. Seemed in a hurry. Mostly close ended questions. Didn't really give her opportunity to be open.
Use of open ended questions	32	Repeatedly asked same questions – no open ended questions – did not pay attention to any answers the child gave – kid kept giving information about his mom's life and it was ignored>. Promised weekly visitation – asked a lot of questions about foster family and their kids but ignored what child said about own life - Didn't follow up w/ child about why wouldn't want to see dad. Don't think took any notes – didn't pay attention to the child's body language – seemed disinterested.
Engagement	31 Sub: wrong language (12), lack of eye contact or problem with body language (15), wrong tone/volume (4)	*Talked so quiet it was hard to hear him. Too many long pauses and silence makes children uncomfortable or you lose their interest. Didn't explain who he was and his reason for being there. Didn't build rapport. No eye contact. Kept head down. Didn't follow up with information given.*
Rapport building	27	
Use of tools	15	"Can you…" just say – tell me… How many vehicles? Talked about truck … transportation? Lack conversation questions and answers. Asked hx DV maybe how they get along. What do children get? Who cooks? Who does grocery shopping? Scale 1-10 how rate self as a mom? Did not ask mom specifically about alcohol use that night or in general. Did not ask about any other drug use or if she ever misused her prescription. I would have asked her to tell about her accident to see if drugs or alcohol were involved. Some areas of the AOCS regarding Adult functioning were not addressed (mental health, domestic violence, support …)
Disclosure	Disclosed incorrect information (3), withheld information (2), made promises (9)	Be careful to not overpromise – did not, but close – regarding visits, etc. Don't be too specific. Don't talk quite so much about staffing w/ "Boss" Promised a visit on Thursday or Friday – don't promise. Explain you role as a worker don't rush this part it is very critical to establish a relationship with the child and confused the child w/ may have another
Prepared for interview	13	Worker did not build rapport with client, gather demographic info, … overview of investigation process, make sure to address all allegations, no … of the condition of the home, did not ask about previous CW involvement….
Provide or explain resources	11	Be careful at telling child they can go to a church w/o checking first with foster parents. Slow down in talking to children. Need to focus more on

		safety in the foster home.
Avoid bias in interview	5	*Seemed more like an interrogation than an interview… Focused more on drug use than anything else…. Seemed intimidating and used a very serious tone through the entire interview…. Said it is concerning that the child walks blocks to school – that's not necessarily too far….. For a 13 y/o. Seemed accusing when pointing out difference in the interviews.*

When comparing the older forms of data collection to those used in the new rubric some specific differences emerge. In the older assessments participants were scored on if they did or did not present with specific skillsets, denoted by a yes or no response. This fails to capture the level of skill these test takers are presenting with, and rather suggests that participants either have or do not have a skill. This fails to capture the experience quantified in the newer rubric. For example, in the new rubric data suggested how workers most often presented with emerging skill competence, even in best-case scenarios. Emergent competence can be difficult to objectively quantify, and could be interpreted as a marginal or borderline display of a practice skill, a majority of those being evaluated using the old system were experiencing what is likely their best possible expected performance being evaluated in a pass or fail manner that could in reality be a very thin and subjective margin between the two. This could explain why the workers appear to be scoring in slightly different ways on the old instruments as opposed to the new ones. The new rubric could help to eliminate this false dichotomy of *competence vs. incompetence* that may have been complicating assessment and worker readiness in the older evaluation materials. Rather, the newer materials, while still capturing the same domains, are able to demonstrate where the workers are in their process of working toward competence and expertise. This type of developmental feedback would be more useful to new workers.

The finalized rubrics that incorporate all of the feedback and comments are located at the end of the report in Appendix A (adult rubric) and Appendix B (child rubric). At the request of DHS, specialized questions were included for both Child Protective Services and Foster Care interviews. The electronic surveys identified if the interviews were for Permanency Planning, Child Protective Services, Foster Care or Adoption. The skip logic utilized in the electronic surveys allowed the specialized questions to be rated only for the appropriate interview type. The finalized rubrics also included some summative items regarding the overall interview and places for qualitative comments related to concerns, strengths and suggestions.

Objective 3: piloting standardized actors and simulation for testing

The HOT process relies on realistic child welfare practice scenarios that are developed to test specific new child welfare worker competencies. Typically, the CW Training staff develops and refines the HOT scenarios at the beginning of each fiscal year. To facilitate this research without adding additional work for already overtaxed CW Trainers, HOT project staff developed 5 permanency planning, 5 child protective services and 1 resource family scenario to be utilized in the simulated interviews. An overall scenario was developed to be given to the actor and interviewer, as well as actor information for both the child and adult interviews and interviewer information for both child and adult interviews. There were a total of five documents for each scenario. The scenarios covered a variety of types of allegations/situations and were all based on real cases. They included parental use of alcohol and drugs, dirty houses and one medical child abuse. The scenarios portrayed situations that on the surface sounded bad and then after interviews were better and vice versa. Also included were interviews with several fathers to encourage new workers to focus on involving fathers.

Although DHS modified some of the scenarios slightly to fit the new worker skill level, the scenarios are still being used for CORE Training and HOT, alleviating the need for CW Training staff to develop new scenarios for Fiscal Year 2015.

Training of Simulated Clients

Since the inception of HOT in 2012, DHS has utilized a variety of persons as "actors" for the simulated interviews, including all levels of training unit personnel, child welfare supervisors and Continuous Quality Improvement staff. Having adequate actors available is an essential component of the HOT process and this has been a constant struggle for the CW Training Unit. One objective of the project was to enhance the simulated interview portion of HOT with the assumption that properly trained professional actors, i.e. simulated clients, would provide a more consistent experience for all participants than just using anyone who is available at the time.

The actors utilized in this project were recruited through OUTulsa Job Application on-line for temporary actors. Those who applied were in a pool to be used for a number of "standardized patient" opportunities. Some had previous acting formal acting experience and some only through the "standardized patient" activities, mainly for the medical and nursing schools. It was not possible to recruit all the correct gender actors for the scenarios, although great efforts were made to change the scripts and scenarios to reflect the actors that were available each day.

Training for all of the HOT actors was provided by Mary Grissom, MSW, from OU CWPEP, with occasional assistance from Allison Holeman, MSW, from DHS. The training:

- Provided copies of the scenarios with the "scripts" for their designated role –adult, child
- Introduced Child Welfare with a brief overview of what Child Welfare does, explaining the different functions within CW (CPS, PP, FC, Adopt, etc.)
- Explained the purpose and importance of HOT and a little about the expected skill level of the interviewers
- Discussed "typical" clients and how they might react. Discussed specific issues – for instance the scenario with the developmentally delayed parent – how that parent might react, talk and understand.
- Discussed the children in custody and different responses that might be seen, tried to help the actors understand DHS custody children and their circumstances
- Explained the role of the "standardized patient" in HOT and that it was important to react to the interviewer, depending on how the "standardized patient" was approached – that is, if the worker was aggressive, it was OK to be aggressive back, not to make it too easy but also not impossible to allow the worker to gather necessary information.

As previously noted, Phase 1 Pilot included adult and child interviews completed by 10 CWPEP practicum students, both BSW and MSW from the Norman and Tulsa campuses. Students from Norman were transported to the Tulsa campus and all interviews were conducted in the SCOUT Simulation Center at OU Tulsa. Phase 2 involved using the simulation center for 55 CORE participants in May and June 2014. Feedback about the simulation experience was obtained from the 1) student pilot group, 20DHS training staff, and 3) simulated actors.

Student feedback on Simulation

Student feedback on the simulation experience included heightened anxiety related to not having seen or role-played the assessment tools prior to this experience; also at having to wait an extended period of time to test, and having only a brief amount of time to both review the case and complete the assessment. They also noted value for this experience, in that it was helpful in providing them the opportunity to try out interaction and rapport-building skills, to receive immediate

feedback from simulated clients and to experience the balance between info-gathering and fostering communication and rapport with clients.

DHS Feedback on Simulation

According to DHS training staff, the HOT pilot events on May 12-13, 2014 and June 24-25, 2014 did not particularly help or hinder the normal DHS process of completing the HOT testing. DHS training staff did all of the pre-planning and scheduling for the flow of the interview days, based on previous experience with HOT and that information was inserted into the SCOUT Simulation Center scheduling process by the SCOUT Special Projects Coordinator. The experience of the DHS and OU schedulers was very evident in that the workers and rooms were coordinated perfectly and the days flowed well. It was very beneficial to DHS that OU provided space at the SCOUT Simulation Center for all of the interviews, as DHS has no facilities for new workers to do HOT testing in Tulsa. In addition, OU provided lunch on-site for all of the workers being tested, the graders and the DHS Training staff, which helped the days go smoothly, and was appreciated by all. It did not appear that the pilot project activities created additional work for any of the DHS staff; in fact, some OU staff members were helpful in guiding new workers throughout the day. The idea that the interviews were going to be videotaped and be part of a research project created additional anxiety for some of the workers being tested, and a small number of workers refused to participate in the study. The process helped DHS see where they could possibly go with simulation if the proper space and equipment was available to them. Also, with written consent from the participants, DHS has utilized some of the interview DVDs as training tools for developing new graders.

Utilization of professional actors as standardized patients in HOT received mixed reviews from DHS staff. Some actors were excellent and others did not work out well. While most actors kept to the script, a few went off script or over-acted, and one, at least, caused some difficulty with a worker/interviewer. However, several workers commented that it was easy to forget that they were actors and it got very "real" quite quickly and acting issues can easily be resolved with training. The benefit of using CW staff rather than actors is that the CW staff members do not require training as they have had actual experience with a wide variety of case situations and are better able to respond appropriately to what the interviewer does and says. CW staff members, however, are often not available for HOT. The benefit of using professional actors would be significant if it were possible to use the same group of actors on a regular basis so that they would gain experience in acting and reacting as a CW parent or child and the best ones could be utilized consistently. It would be helpful to recruit some teens to do both the teen scenarios and those of younger children.

Recommendations

The following are recommendations based on the multiple sources of data and feedback on the process of piloting simulation for the HOT test.

✓ Standardize the HOT testing process

Since the inception of HOT in 2012, DHS has utilized a variety of persons as "actors" for the simulated interviews. There has not been a standardized system used for the training of these clients or a widespread agreement on how the characters should behave in interviews, other than the practice wisdom and experience of the volunteer clients. Sometimes the clients have been janitors, receptionists or other staff with little to no child welfare experience. This lack of

standardization is limiting in several ways: (1) it reduces the objective nature of the assessment, increasing the potential of the delivery of an unfair test, (2) there are few opportunities to develop a more nuanced and systematic assessment of workers, (3) collecting data 'old school' via paper and pencil evaluation allows for a more difficult analysis of data trends in new workers over time, (4) there are fewer opportunities to provide workers with multi layered improvement data. Using the simulation center, trained actors and standardized scenarios allowed the research team to explore the potential and possibility of using this mechanism for testing into the future. The following explores these points further.

✓ Utilize the rubrics developed through this project

Based on the analysis conducted through this project, the adult and child interview rubrics' inter-rater reliability is very high. Each item in the rubric has been operationalized in descriptive, behavioral terms and is based on review of literature pertaining to key interaction and interviewing skills. Such operationalization allowed for clarity and consistency in raters' understanding of skills being tested, and contributed to the high inter-rater reliability demonstrated. The new rubric may be able to alleviate some of the concern voiced by CW Supervisors and others that the existing HOT process is not a fair or valid way to evaluate new workers because of the level of subjectivity in the grading process.

Unlike the current DHS grading forms that allow only for a nominal level of assessment (yes/no responses to demonstrated worker skills), the rubrics utilize a 0-4 Likert-type scale for rating employees' mastery of various skills on a more nuanced and descriptive scale. This structure also allows raters to better locate testers in terms of the developmental process of the skills tested. Rubric ratings range from *not competent* through *emerging competence* to *very competent*, rather than just a yes/no binary indication of skill. Analysis of rubric ratings can provide information as to how well individual employees and new workers as a whole are doing on various competencies and skills and can be used to develop individual development plans, to analyze trainers' effectiveness and to evaluate and refine the CORE curriculum. A pass/fail score can also be developed from an analysis of the rubric data. It is also recommended that when rubrics are utilized, that they are incorporated into the assessment process in one single form for a period of time (1-3 years) so that comparative data is available for meaningful comparison and feedback. Frequent changes to instrumentation complicate evaluative processes, and often make it more difficult to extract meaningful data to improve programming. The final revised rubrics that incorporate all feedback from raters are available in this report in Appendices A&B.

✓ Structure HOT test interviewing within a developmental framework.

One conclusion of the research team was that there are actually distinct skill sets being tested, that these skills are developmental in nature (i.e. involve a process of learning to engage and build rapport with clients as well as a separate process of learning to gather information with which to assess safety), and that as a result, testing should reflect this process. It is recommended that communication, rapport-building, and essential engagement skills be taught and tested as a distinct set of skills that transcend one area of child welfare practice. Information-gathering and safety assessment skills are distinct, yet related, but also differ during testing based upon the worker's unit. Variance between these work categories (permanency planning, etc.) dictates difference goals and methods of information-gathering, and difference decision-making processes based on the client interaction; therefore, it is recommended that tools assessing information-gathering and safety assessment also be tailored to the job and separated from the assessment of essential interaction skills. As noted, these are separate skill sets, and one must master interaction skills before applying them to a high-stakes

interaction intended to assess safety. In terms of testing procedures, this points to a two-tiered process of simulated interviews, rather than one lengthy interview which includes both.

If possible, we recommend that the interaction skills be taught and tested for all workers in CORE , while information-gathering be incorporated in Level I and II Permanency Planning, CPS, Foster Care and Adoption training. Another possibility is to do the child interview to demonstrate engagement skills and the adult interview to demonstrate assessment skills utilizing specific assessment tools or vice versa. This would allow both testers and raters to focus on one area of communication skills, shorten each session to a more manageable length for raters, and provide targeted feedback for skill-building.

Additionally, a finding here revealed that safety assessment skills were the lowest rated skills set for testers using the new rubric. Safety assessment was not well captured in the existing HOT evaluation tools. Since assessment of safety is a critically important element of child welfare work, this area can be strengthened with specialized training and skill evaluation in which the ability to determine safety is the main or only focus. Lastly, by understanding workers' capabilities and skills in a developmental framework would help ensure that raters have realistic expectations of new workers as those skills related to nuanced categories ranging from not competent, emerging competence, etc.

✓ Undertake the difficult and time consuming task of developing competencies based on what DHS believes new CW workers need to have in regard to knowledge, skills and abilities

OKDHS can consider the development of competencies for workers as they progress to Level I and II and then develop training that teaches to the competencies and evaluate HOT participants based on those competencies. From this project as well as previous work on the Level I and II certification process, it appears as though current competencies have been based on what content is included in an existing workshop rather than what the agency has determined the worker needs to know or be able to do at various stages of professional development.

✓ Create more opportunities for observation, practice interactions, and simulated interviews prior to the actual high stakes testing of HOT.

It is recommended that efforts be made to incorporate simulation activities into CORE training as frequently as possible to help workers practice interviewing skills and develop confidence for the HOT interviews. Such practice allows for the new workers to practice skills and training knowledge in a safe environment that is not as high stakes as either the final HOT test or a real interview with a child or parent. Use of role-plays, interaction analysis, process recordings, and live simulations are all strategies that would allow for multiple experiences with key skill areas, the incorporation of feedback from several sources, and for a developmental process to unfold over time. This developmental process also highlights the need for not only summative evaluation of skills (one time, after training) but formative evaluation, or evaluation of skill-development from basic to advanced as a process. Specifically:

- During CORE, provide participants with opportunities to complete practice interviews with a well-trained simulated client rather than just role playing with another participant. Video tape the simulated practice interviews for workers to take with them to review and analyze, share with their supervisors and identify areas for improvement prior to HOT.

- In support of the new DHS Child Welfare Core Strategy #1 Hiring and Retention of Staff, utilize CQI staff who are assigned to various regions to work directly with new workers right after CORE but prior to HOT, providing opportunities for practice interviews with immediate feedback and establishing a development plan with the worker, identifying specific things they need to do to prepare for HOT or the retake.

✓ Train the graders of the HOT test in a systematic manner.

Develop training for individuals responsible for grading the HOT test. Provide the rubric, clarify the categories and the purpose of the HOT test. If adopted, talk to graders about the developmental progression of worker skills and how to provide both formative and summative feedback. Provide specific examples, either through written comments or short videos that clarify the difference between competence versus emerging competence or poor skill on specific behaviors.

✓ Utilize electronic data collection.

Regardless of whether the new rubrics or some other rating forms are utilized, develop an electronic survey to gather the rating data on an ongoing basis. Use of an electronic tool is recommended, as this option allows for large-scale and longitudinal data management where data is recorded at the time of testing, in a consistent manner, rather than time-consuming data entry after the fact. Additionally, electronic data management allows for manipulation of information to assess not only an individual tester, but testers across time and place. With this ability, trends in knowledge and skill deficits and strengths may become more apparent, and training efforts informed by findings.

Another possibility is to use portable tablets that graders can utilize during interviews. If such tablets are not possible, let graders use paper forms, but require that they immediately enter the ratings into the database before they leave for the day. This will make the data quickly available for review and analysis.

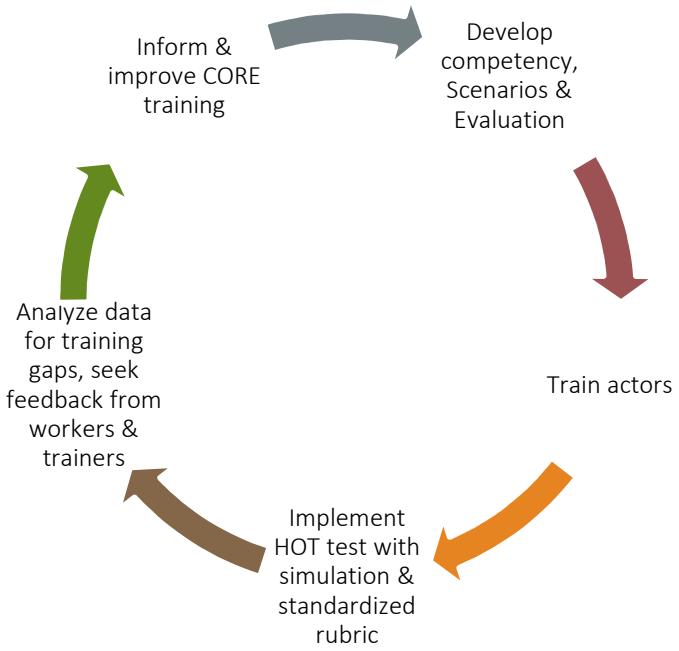
Finally, in the "Cadillac version" of electronic data collection, electronic surveys and video annotations (available as part of the simulation software) that can incorporate instructor, actor, and self-rated feedback would allow for the collection and use of data at both the individual and aggregate level. Such data would allow for a more refined use of data that can be used by the individual workers to focus developmental improvement and for the DHS staff in terms of training. This is elaborated further in the final recommendation below.

✓ Consider using the simulation center software, electronic data collection and standardized actors for the HOT test.

There are several advantages to considering using the full spectrum of simulation, standardized actors and electronic data collection through the OU simulation center (Or a similar facility) for the HOT testing format. There is emerging research evidence that the use of simulation for the training and assessment of child welfare workers may be an effective method for improving and rating worker skills (Lee, Stapes, Mankowski, 2015; Friend, 2009) and in the training of social work skills (Badger & MacNeil, 2002; Logie et al, 2013). When simulations are rooted in real-life situations, including realistic depictions of clients and subject matter, they are more likely to experience skill transfer. These simulated environments are far more than traditional role-plays, as they create more realistic experiences of client interaction. The following further explores the advantages of this recommendation:

- The reduction of the subjective nature of the assessment of new CW workers during the high stakes HOT test. Standardized patients are individuals who are specially trained to act as clients or patients for purposes of instruction, practice and assessment of the clinical or examination skills of a wide variety disciplines. They are a significant resource for teaching and evaluating clinical skills, including: interviewing skills, assessment, physical examination, communication, history taking, patient counseling and patient education. Professional trainers, staff, and faculty members prepare standardized patients for their roles in all assigned cases. Standardized patients are able to assist in both teaching and evaluation of students. They also are trained to provide constructive, non-threatening feedback regarding the completeness and style of an encounter. These standardized patients can be used one on one, in small or large groups, family sessions or other sessions. Using standardized clients for child welfare testing takes advantage of this and ensures that the actors understand the spectrum of behaviors and responses that a person in a given situation might respond. They can be coached to respond in particular ways, depending on the skills of the student/worker. This ensures that particular events happen in a particular order in an interview, depending on the skills of the individual being tested. Much less is left to chance with trained actors thus reducing the potential subjectivity of the test itself.

- The provision of feedback to individual workers that focuses on specific strengths and areas for improvement. The simulation center offers the opportunity to take advantage of technology for the collection of data. For example, workers could perform test at the simulation center. The evaluator could access the video online (even offsite or in a different room), annotate the video with specific feedback linked to specific moments in the interview, conduct the evaluation as they are reviewing the video (including making comments), and submit the evaluation. The worker could view the video and evaluation immediately after the interview. In addition, the standardized patient could also rate the worker (on select items related to rapport, communication, etc) and this data would be immediately entered by the client post interview. All of these data, including videos, would be captured electronically and could be analyzed for trends. If for example, workers repeatedly had a difficult time with a particular scale; training could be adapted to address this skill. Further, even workers that pass the HOT test would benefit from an analysis of key strengths and areas for skill improvement. Workers would also be provided with a link to the video and a self-assessment. They would be able to review the video with annotations, as well as view evaluator and client feedback. The video and evaluation data could also be shared with the worker's supervisor and this data could be used to help develop an improvement plan between the worker and supervisor.



- The provision of data from multiple sources that provide specific and detailed feedback that improves CORE training. The use of data from multiple sources (rubrics, self-rating, actors, video annotations) will allow for more detailed feedback on aggregate worker performance that can inform the CORE training and help to further articulate the most important skills and competencies needed for a beginning worker at OK DHS.

Summary

This project involved the development and testing of standardized rubrics and simulation for the testing of new child welfare workers in OKDHS. Findings reveal a new measure that is reliable and valid and overall positive feedback about the potential to use simulation for HOT testing in the future. A list of recommendations are offered that highlight the major findings of the project.

References

Badger, L., & MacNeil, G. (2002). Standardized clients in the classroom: A novel instructional technique for social work educators. *Research on Social Work Practice, 12*(3), 364-374.

Barry, A. E., Chaney, E. H., Stellefson, M. L., & Chaney, J. D. (2011). So you want to develop a survey: Practical recommendations for scale development. *American Journal of Health Studies, 26*(1), 97-105.

Carpetto, G. (2008). *Interviewing and brief therapy strategies: An integrative approach*. Boston, MA: Pearson Education

Carrillo, D. F., Gallant, P. Thyer, B. A. (1993). Training MSW students in interviewing skills: An empirical assessment. *Arete, 18* (1), 12-19.

Cheung, M. (2008). Promoting effective interviewing of sexually abused children: A pilot study. *Research on Social Work Practice, 18* (137), 137-143. DOI: 10.1177/1049731507304359

Friend, C. (2009). Helping child welfare workers learn interviewing skills: A research report. Berkeley: University of California at Berkeley, California Social Work Education Center.

George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide andreference. 11.0 update (4th ed.). Boston: Allyn & Bacon.

Lee, H., Staples, H. & Mankowski, H (2015, Feb). MIND THE GAP NATIONAL WEBINAR SERIES, Session #2 Simulation Labs for Child Welfare Training & Education https://ncwwi.org/files/NCWWI_MindtheGap_SimulationLabs.pdf

Logie, C., Bogo, M., Regehr, C., & Regehr, G. (2013). A critical appraisal of the use of standardized client simulations in social work education. *Journal of Social Work Education 49*(1), 66-80.

Pike, C. K., Bennett, R., & Chang, V. (2004). Evaluation of two interviewing skills measures: An instrument validation study. *Advances in Social Work, 5* (1), 61-75.

Stevenson, K.M., Leung, P., & Cheung, K.M. (1992). Competency-based evaluation of interviewing skills in child sexual abuse cases. *Social Work Research and Abstracts, 28* (3), 11-16.

Zastrow, C. (n.d.) Interviewer skills rating sheet. Fitchburg, WI: Author.

Final Rubrics

Both the adult and child rubric use the following scale to assess performance on the HOT test.

Scale:	
0	Poor, no evidence of skill, not competent
1	Fair, lacks clear evidence of skill, limited competence
2	Good, some evidence of skill, emerging competence
3	Very good, clear evidence of skill, competent
4	Excellent, ample evidence of skill, very competent

Appendix A: HOT Testing Rubric Adult– Revised Final Version

Rapport-Building Skills: please evaluate participant's rapport-building skills on the scale of 0-4 seen below.	0	1	2	3	4
1. **Introductions** - clearly identified self and role; addressed client appropriately; used names	❍	❍	❍	❍	❍
2. Established **clear expectations**; explained purpose and process of interview; explained confidentiality	❍	❍	❍	❍	❍
3. **Respectful, nonjudgmental stance** (refrained from blaming; active listening, etc.)	❍	❍	❍	❍	❍
4. Demonstrated **empathy** (able to accurately reflect feelings of client; indicated understanding and interest; expressed concern; acknowledged client's emotions)	❍	❍	❍	❍	❍
5. **Non-threatening**, approachable demeanor (receptive, open posture, eye contact, warmth/welcoming voice tone, etc.)	❍	❍	❍	❍	❍
6. **Professionalism** (courteous; appropriate presentation; use of professional language; boundaries; etc.)	❍	❍	❍	❍	❍
7. **Strengths-based** attitude (highlighted client abilities and strengths in addition to concerns)	❍	❍	❍	❍	❍
8. Ability to use age-**appropriate language** and to adjust interaction to client (due to age, mental or emotional state, etc.)	❍	❍	❍	❍	❍

Communication and Information-Gathering Skills: please evaluate participant's communication and information-gathering skills on the scale of 0-4 seen below.	0	1	2	3	4
1. **Used focused, open ended questions** (vs. yes/no or asking more than one question at a time)	❍	❍	❍	❍	❍
2. **Use of silence** - allowed client to speak, gather thoughts, did not interrupt	❍	❍	❍	❍	❍
3. **Active listening** - eye contact, neutral responses (head nod, attentive posture, etc.)	❍	❍	❍	❍	❍
4. Appropriate use of **paraphrase or summarization** to reflect content of client statements	❍	❍	❍	❍	❍
5. Use of **developmentally-appropriate terms**; refrain from use of jargon; explained unfamiliar terms	❍	❍	❍	❍	❍
6. **Ability to join with client** by using language mirroring and client-defined terms	❍	❍	❍	❍	❍
7. **Able to redirect client** and/or keep interview focused on its stated purpose	❍	❍	❍	❍	❍
8. Employed **probing and exploration** questions to gather information	❍	❍	❍	❍	❍
9. **Avoided leading and coercive questions** (gives choices, does not suggest one response over another, etc.)	❍	❍	❍	❍	❍
10. **Allowed client to ask questions**/provided appropriate responses to question	❍	❍	❍	❍	❍
11. Demonstrated ability to **question inconsistencies** or confront in a respectful manner	❍	❍	❍	❍	❍
12. **Checked in with client** during interview to assess for understanding	❍	❍	❍	❍	❍

Safety Assessment and Ending Skills: please evaluate participant's safety assessment and ending skills on the scale of 0-4 seen below.	0	1	2	3	4
1. Gathered necessary information for **assessing client safety**	❍	❍	❍	❍	❍
2. Explored safety issue from **client perspective**	❍	❍	❍	❍	❍
3. Gauged client's **understanding of safely** concern(s) before terminating interview - for example, asked for client summary of critical information; rephrased if lack of client understanding is evident	❍	❍	❍	❍	❍
4. Used **probing questions** to gather further information about safety (context and nature of concern)	❍	❍	❍	❍	❍
5. **Validated client** experiences and concerns - received information with care and concern	❍	❍	❍	❍	❍
6. **Follow up plan**/next steps made clear	❍	❍	❍	❍	❍
7. **Refrained from making promises** or providing premature predictions about the future	❍	❍	❍	❍	❍

NOTE→ ONLY USE THIS IF PARTICIPANT'S CASE TYPE WAS CHILD PROTECTION SERVICES! Otherwise, skip entirely.	0	1	2	3	4
Worker assured client that they were not in trouble with the worker	❍	❍	❍	❍	❍
Worker addressed the allegations	❍	❍	❍	❍	❍
Worker asked questions to determine the client's understanding of the reason for CW concern	❍	❍	❍	❍	❍

NOTE → ONLY USE THIS IF PARTICIPANT'S CASE TYPE WAS FOSTER CARE! Otherwise, skip entirely.	0	1	2	3	4
Worker asked questions to gather information regarding the client's perception of foster care and being a Bridge Family	❍	❍	❍	❍	❍

Overall, please rate your impression of participant's:	0	1	2	3	4
pacing of the interview	❍	❍	❍	❍	❍
ability to focus on the interaction and remain sensitive to the client	❍	❍	❍	❍	❍
readiness for the field	❍	❍	❍	❍	❍

In the space below, please give details or examples to illustrate concerns.

In the space below, please give details or examples to illustrate strengths.

In the space below, please give details or examples to illustrate suggestions.

Appendix B: HOT Testing Rubric Child – Revised Final Version

Scale:

0	Poor, no evidence of skill, not competent
1	Fair, lacks clear evidence of skill, limited competence
2	Good, some evidence of skill, emerging competence
3	Very good, clear evidence of skill, competent
4	Excellent, ample evidence of skill, very competent

Rapport-Building Skills: please evaluate participant's rapport-building skills on the scale of 0-4	0	1	2	3	4
1. **Introductions** - clearly identified self and role; addressed client appropriately; used names	❍	❍	❍	❍	❍
2. Established **clear expectations**; explained purpose and process of interview; explained confidentiality	❍	❍	❍	❍	❍
3. **Respectful, nonjudgmental stance** (refrained from blaming; active listening, etc.)	❍	❍	❍	❍	❍
4. Demonstrated **empathy** (able to accurately reflect feelings of client; indicated understanding and interest; expressed concern; acknowledged client's emotions)	❍	❍	❍	❍	❍
5. **Non-threatening**, approachable demeanor (receptive, open posture, eye contact, warmth/welcoming voice tone, etc.)	❍	❍	❍	❍	❍
6. **Professionalism** (courteous; appropriate presentation; use of professional language; boundaries; etc.)	❍	❍	❍	❍	❍
7. **Strengths-based** attitude (highlighted client abilities and strengths in addition to concerns)	❍	❍	❍	❍	❍
8. Ability to use age-**appropriate language** and to adjust interaction to client (due to age, mental or emotional state, etc.)	❍	❍	❍	❍	❍

Communication and Information-Gathering Skills: please evaluate participant's communication and information-gathering skills on the scale of 0-4 seen below.	0	1	2	3	4
1. **Used focused, open ended questions** (vs. yes/no or asking more than one question at a time)	❍	❍	❍	❍	❍
2. **Use of silence** - allowed client to speak, gather thoughts, did not interrupt	❍	❍	❍	❍	❍
3. **Active listening** - eye contact, neutral responses (head nod, attentive posture, etc.)	❍	❍	❍	❍	❍
4. Appropriate use of **paraphrase or summarization** to reflect content of client statements	❍	❍	❍	❍	❍
5. Use of **developmentally-appropriate terms**; refrain from use of jargon; explained unfamiliar terms	❍	❍	❍	❍	❍
6. **Ability to join with client** by using language mirroring and client-defined terms	❍	❍	❍	❍	❍
7. **Able to redirect client** and/or keep interview focused on its stated purpose	❍	❍	❍	❍	❍
8. Employed **probing and exploration** questions to gather information	❍	❍	❍	❍	❍
9. **Avoided leading and coercive questions** (gives choices, does not suggest one response over another, etc.)	❍	❍	❍	❍	❍
10. **Allowed client to ask questions**/provided appropriate responses to question	❍	❍	❍	❍	❍
11. Demonstrated ability to **question inconsistencies** or confront in a respectful manner	❍	❍	❍	❍	❍
12. **Checked in with client** during interview to assess for understanding	❍	❍	❍	❍	❍

Safety Assessment and Ending Skills: please evaluate participant's safety assessment and ending skills on the scale of 0-4 seen below.	0	1	2	3	4
1. Gathered necessary information for **assessing child safety**	❍	❍	❍	❍	❍
2. Explored safety issue from **client perspective**	❍	❍	❍	❍	❍
3. Gauged child's **understanding of safely** concern(s) before terminating interview - for example, asked for client summary of critical information; rephrased if lack of client understanding is evident	❍	❍	❍	❍	❍
4. Used **probing questions** to gather further information about safety (context and nature of concern)	❍	❍	❍	❍	❍
5. **Validated client** experiences and concerns - received information with care and concern	❍	❍	❍	❍	❍
6. **Follow up plan**/next steps made clear	❍	❍	❍	❍	❍
7. **Refrained from making promises** or providing premature predictions about the future	❍	❍	❍	❍	❍

NOTE → ONLY USE THIS IF PARTICIPANT'S CASE TYPE WAS CHILD PROTECTION SERVICES! Otherwise, skip entirely.	0	1	2	3	4
Worker assured child that they were not in trouble with the worker	❍	❍	❍	❍	❍
Worker addressed the allegations	❍	❍	❍	❍	❍
Worker asked questions to determine the client's understanding of the reason for CW concern	❍	❍	❍	❍	❍

NOTE → ONLY USE THIS IF PARTICIPANT'S CASE TYPE WAS FOSTER CARE! Otherwise, skip entirely.	0	1	2	3	4
Worker asked questions to gather information regarding the child's perception of foster care and being a Bridge Family	❍	❍	❍	❍	❍

Overall, please rate your impression of participant's:	0	1	2	3	4
pacing of the interview	❍	❍	❍	❍	❍
ability to focus on the interaction and remain sensitive to the client	❍	❍	❍	❍	❍
readiness for the field	❍	❍	❍	❍	❍

In the space below, please give details or examples to illustrate concerns.

In the space below, please give details or examples to illustrate strengths.

In the space below, please give details or examples to illustrate suggestions.