ILLINOIS

CFRC | Children & Family Research Center

SCHOOL OF SOCIAL WORK

FY2021 Program Evaluation of the Child Protection Training Academy for New DCFS Investigators

Theodore P. Cross
Yu-Ling Chiu
Shufen Wang
Laura Lee
Steve Tran
Kirsten Havig

June 2021

Prepared for:
Illinois Department of Children and Family Services
Marc D. Smith, Acting Director

Table of Contents

Executive Summary

Since the Child Protection Training Academy (CPTA) launched the first simulation training at University of Illinois at Springfield (UIS) in February 2016, the CPTA has trained over a thousand new child protection investigators hired by the Illinois Department of Children and Family Services (DCFS). Trainees receive first-hand experience learning a wide range of child protection tasks, from the first knock on a family's door to testifying in family court, guided by expert trainers and working with actors playing the family in a mock house and mock courtroom. In FY2021, the Children and Family Research Center's (CFRC) evaluation team again used multiple sub-studies to examine the implementation and outcomes of simulation training. This is an important time historically to study simulation training because of the effect of COVID-19 on trainees, their work and the training itself.

Chapter 1: Introduction

The COVID-19 pandemic had the country in its grip in FY2021 and the Child Protection Training Academy had to adapt to it. After a break of five months, the CPTA began virtual simulation training of new child protection investigators in both the Chicago and Springfield laboratories using Zoom. In this report, we present results from trainings that used virtual methods. Thus we can see if trainees' reactions to the virtual simulations differed from their reaction to the in-person simulations.

The CPTA began to train a new group of DCFS professionals this fiscal year: supervisors of investigators. The training team used a number of experiential methods involving exposing supervisors to visual or textual representations of families, and training the supervisors to use problem-based learning (PBL) to guide their inquiry and decision-making about these families. The methods for evaluating problem-based learning are in development. We anticipate adding a component on the evaluation of this training in the future.

In addition to doing simulation training online, the Child Protection Training Academy changed the format of simulation training for investigators during the latter part of this fiscal year. CPTA spread simulation training across parts of two weeks rather than concentrating it one week. CPTA also re-formatted investigator training to include two case scenarios rather than just one. This provided trainees a broader experience and gave them extra practice in family engagement.

Chapter 1 also provides an overview of previous program evaluations results between FY2017 and FY2020. Employing a range of different quantitative substudies, program evaluation has detailed the implementation of simulation training; assessed trainees' reactions to the training during the training, immediately after and after a follow-up period; and examined the relationship of training to employee turnover.

Chapter 2: Daily Experience of Simulation Training (DEST)

The Daily Experience of Simulation Training (DEST) measure was designed to examine change in trainees' confidence over the course of simulation training. During the week of simulation training, trainees rated their confidence daily on 13 child protection work skills. This is an important time to assess DEST results, because of changes in simulation training during FY2021,

as discussed in the Introduction. A new version of the DEST (DEST 3.0) with additional questions concerning trainees' on-the-job training and the effectiveness of the debriefing was implemented in FY2021.

Between August 24, 2020, and April 28, 2021, a total of 138 trainees participated in the simulation training. All of them filled out the DEST at one time point or more during the training. The DEST data included 737 responses across six time points. The weighted average daily response rate was 89%. Moreover, out of 138 respondents, 84 (61%) completed the DEST at all six time points.

All 13 confidence items showed a substantial linear increase over the course of simulation week. The average trainee's confidence level increased steadily from baseline to the last day across all 13 items. One-way ANOVAs with linear contrasts were statistically significant, indicating a significant linear increase in confidence over the course of the simulation-training week for all 13 skills. A repeated measures ANOVA was conducted with the 84 respondents who completed the DEST at every time point. Consistent with the findings above, the confidence of respondents on performing the 13 investigative skills showed a significant linear increase over the course of simulation training week and the effect sizes were in the medium to large range (i.e., $\eta2 = .05$ to $.11$, or $d = .64$ to 1.10). Comparing DEST results across training cohorts enables us to see if changes in trainees' confidence have been consistent across trainings. The results suggest that most cohorts, on average, experienced meaningful increases in confidence during virtual simulation training that resemble the increases respondents in the previous year experienced with in-person training.

Due to the pandemic, some trainees might have experienced a greater delay in receiving the simulation training. Consequently, they may have received more on-the-job training (OJT) prior to receiving simulation training than in previous years. Analysis with Kendall's tau statistic showed a small but statistically significant relationship between OJT and confidence. Those with longer OJT tended to be slightly more confident than those with shorter OJT, at baseline ($r=.18$, $p<.05$), Day 1 ($r=.19$, $p<.01$), Day 2 ($r=.16$, $p<.05$), and Day 4 ($r=.24$, $p<.01$).

CPTA changed the training format twice this fiscal year, as mentioned earlier. In response to the COVID-19 pandemic, CPTA delivered the training online. In addition, in February 2021, CPTA changed the timeframe to Thursday to Wednesday and added an additional mock case. Analysis of variance tested the differences among three time periods: prior to August 2020 (pre-virtual), between August 2020 and February 17, 2021 (virtual), and between February 18 and April 28, 2021 (virtual and new timeframe). For both Day 2 and Day 3, there was no significant difference in respondents' confidence in the 13 skills among the three timelines.

In the DEST, we asked participants to rate the helpfulness of the training team's feedback and the effectiveness of individual and group debriefings. Most respondents found the feedback during simulation training either very helpful or helpful. Respondents were also asked to rate the effectiveness of debriefing. The analysis showed that the group debriefings on Day 2 (knock on the door simulation), Day 3 (scene investigation simulation), and Day 4 (individual interview

and medical simulations) more effective than the group debriefing on Day 1 (calling the reporter simulation). The average ratings of individual debriefing on Day 2 were higher than those on Day 3.

Chapter 3: Post-Training Satisfaction Survey

All new investigators were invited to complete a post-training satisfaction survey about their classroom training and simulation training. This chapter reports trainees' satisfaction ratings for simulation training over this time period. It also compares results for simulation training, classroom training and for the training program as a whole. The chapter also provides qualitative results from the analysis of open-ended items in the post-training satisfaction survey. The post-training survey includes 27 questions about classroom training, 8 questions about simulation training, and 2 questions about the overall training program. There were 50 survey respondents between February 2020 and February 2021 that were included in the analysis for this chapter.

For most of these scales of simulation training satisfaction, the average score was between "undecided," and "agree." This indicated some degree of satisfaction with simulation training on average. We compared mean satisfaction ratings for the current fiscal year to mean satisfaction ratings from previous fiscal years. The trainees' responses from this year ranged from "undecided" to "agree" on eight scales measuring simulation training; whereas, the responses from the previous years ranged from "agree" to "strongly agree" on the same eight scales. The trainees' satisfaction scores for the classroom and overall program averaged between "agree" and "strongly agree"; while the average satisfaction score for simulation training averaged between "undecided" and "agree" and was significantly lower compared to classroom or overall satisfaction. Most of the differences between this year and previous years stemmed from two training cohorts in which several respondents reported a challenging experience (see below) and had substantially lower scores on the satisfaction items.

The online post-training satisfaction survey includes open-ended items in which trainees can write comments in text. We combined respondents' answers to the two open-ended questions and analyzed the text from both questions together. Altogether 34 trainees provided comments about simulation training. Results of thematic analysis showed that many comments focused on the benefits of simulation training while others described challenging experiences in the simulation laboratory. Comments focusing on the benefits were distributed across eight different trainings. The benefits mentioned were:

- The importance of application of classroom information to enacted scenarios
- The value of learning about initial engagement with the family
- The opportunity to watch all the debriefing sessions
- The experience of entering the home and seeing the children and home environment
- The realism of the experience (even when done online)
- The opportunity to learn to deal with difficult situations

- The emphasis on the importance of learning one's cases
- The value of receiving a recording of the debriefing to process feedback at a later time
- The importance of a safe learning environment
- The value of helpful and encouraging feedback
- The value of the help, respect, and validation of the simulation trainers.

Comments that described simulation training as challenging came from four different trainings—all but two of the comments that described significant challenges came from two trainings. Themes that emerged among the group that had challenging experiences were: feeling disrespected, feeling that some information they were presented with in simulation training conflicted with their classroom training, and feeling that one actor was excessively aggressive. Other comments expressed concern related to communication, attention, organization, being provided with adequate information, and cultural competence. Some respondents specifically mentioned the limitations of doing simulation training virtually rather than in-person.

Chapter 4: Turnover Study

This chapter examines whether simulation training is related to employee turnover. Using two different analytic methods, it asks whether investigators trained using simulation training have stayed in their jobs longer than investigators who were not provided simulation training. This inquiry updates the CFRC's previous analysis of the relationship between simulation training and turnover. In its FY2019 report, the program evaluation team used the statistical method of survival analysis to compare investigators who had received simulation training (the sim group) to investigators who had not received simulation training (the non-sim group) on the likelihood of leaving their job. In the current analysis, we again compared investigators who started to work at the Division of Child Protection (DCP) after February 2016 (this group received simulation training) to investigators who started to work at DCP before February 2016 (this group did not receive simulation training). We obtained an updated data file with employment data from DCFS. This file differed from the parallel data file we received from DCFS in FY2019. The new data set included more recent data, of course, but it was also a more complete data set for earlier periods as well. The caveat applies that a comparison of sim and non-sim investigators is confounded with era and we cannot fully resolve the ambiguity this introduces.

One method we used to compare sim-trained and non-sim-trained was interrupted time-series analysis (ITS). This involves looking at the trend in turnover rates over time for investigators and seeing if the introduction of simulation training changed ("interrupted") the trend in turnover rates. One advantage of ITS is that we can see overall trends over time in turnover in addition to the comparison between investigators trained in the sim training era and those trained before that. For this analysis, we defined turnover as leaving an investigator position in the Division of Child Protection (DCP) within 12 months of starting it. The results of the interrupted time series regression analysis indicated that there was a significant drop in the one-year turnover rate following the introduction of simulation training followed by an increase over

time. This current ITS study found that turnover increased for those investigators starting in 2014, went down for those investigators starting in 2016, and went back up again for those investigators starting in 2019. There could be historical reasons for these changes – for example, perhaps COVID-19 influenced the increase in turnover for those investigators starting in 2019. The turnover rate for those who had simulation training is less than for those who did not have simulation training for the period of time studied. However, there is no way to be sure whether this difference is due to simulation training versus other historical events that were contemporaneous with the introduction of simulation training.

As we did in FY2019, we also conducted a survival analysis to compare the sim group and non-sim group on turnover. Survival analysis yields the probabilities that staff will leave the investigator positions during a given time interval. We added an additional feature to the current survival analysis that we did not use in FY2019. To control for differences between sim and non-sim investigators on such variables as experience and age, we created samples in which each sim investigator was matched with a comparable non-sim investigator. This created matched samples that were similar on almost every variable. We used a method called propensity score matching (PSM). The life table of survival analysis showed that through the first 6 months on the job, sim-trained investigators were somewhat less likely to leave their job. From the ninth to the twelfth month on the job, however, sim trained investigators were more likely to leave their job. It should be noted that the first 12 months for the sim investigators and the first 12 months for the non-sim investigators occurred in different historical eras. Note that the time period for the sim investigators includes time in which the COVID-19 pandemic affected the country. The results from the Cox regression the odds of the sim group leaving their job within 12 months were 1.50 greater than the odds of the non-sim group doing so. The results from the current survival analysis differ substantially from the survival analysis we conducted in FY2019. While in FY2019 report we found that sim-trained investigators were less likely to leave their jobs within the observation period, in the current analysis we found that sim-trained investigators were more likely to leave their jobs within the observation period. While several differences between the two analyses might contribute to the difference in findings, the most likely explanation is that the current analysis includes investigators who started at DCP in calendar year 2019, many within 12 months of the onset of the COVID-19 pandemic. We think the difference in the historical eras studied in the two analyses is likely to be the biggest explanation for the difference in results.

Chapter 5: Maltreatment Re-report Study

One ultimate goal of the CPTA's simulation training program is to increase children's safety. To evaluate the relationship of simulation training to child safety, we compared sim-trained and non-sim-trained investigators on the likelihood that children in their investigations were involved in re-reports to DCFS.

We made two comparisons to examine the effect of simulation training on maltreatment re-reports. The first comparison contrasted sim-trained and non-sim-trained on investigations

they worked on during the first two years in their career as an investigator. Thus, in this comparison, the investigators had similar levels of experience. We call this the Equivalent Experience Comparison. This comparison had the disadvantage that sim and non-sim investigations being compared came from different historical periods (because all of the investigators hired after February 2016 had sim training, while none of the investigators hired before that had it). The second comparison contrasted sim-trained and non-sim-trained investigators on investigations during the same historical time period—we call this the Equivalent Historical Period Comparison. The second comparison had the disadvantage of comparing sim- and non-sim investigators with different levels of experience. Because sim-trained investigators were hired later, historically, than non-sim-trained investigators, they have less experience than non-sim-trained investigators in any given historical period. In the Equivalent Historical Period Comparison, sim-trained investigators had 0 to 2 years of experience in DCP while non-sim-trained investigators had 3 to 4 years of experience.

Cox's proportional hazard regression, a type of event history analysis, was used to assess the relationship of investigator simulation training to a substantiated maltreatment re-report in the investigator's cases. For each investigation, the observation period started on the date at which there was a finding in the investigation (the finding date) and continued until either a substantiated maltreatment report occurred, or the 6-month follow-up period ended. Because the unit of analysis was the investigation and each investigator could conduct multiple investigations, it was necessary to adjust the analyses to account for the fact that observations were not independent of one another. Therefore, the Cox's regression procedure was run with the COVSANDWICH variance adjustment, which uses the robust test to evaluate the effects of the predictor variables in the model.

For the Equivalent Experience Comparison, there were 270 investigators in the simulation group and 134 investigators in the non-simulation group. The simulation group conducted 49,451 investigations while the non-sim group conducted 23,637 investigations. The hazard ratio (HR) for simulation training was 1.32, a statistically significant but small effect that was below Azuero's threshold for clinical relevance. A re-report was more likely when a sim investigator investigated the original allegation, though the difference was small. Note that there was also a small but statistically significant relationship of year of original investigation to re-report. Re-reports were more likely in the later years represented in the sample than in the earlier year. Because all the original investigations in the Equivalent Experience Comparison sample in 2014 and 2015 were investigated by non-sim investigators, and all the cases in 2018 thru 2020 were investigated by sim-trained investigators, year is confounded with the sim vs. non-sim comparison, and it is impossible to disentangle the two variables fully. It is impossible to rule out year of investigation as an explanation for the difference between the sim and non-sim groups on re-report.

For the Equivalent Historical Period Comparison, there were 270 investigators in the simulation group and 88 investigators in the non-simulation group. The sim group conducted 49,451

investigations while non-sim group conducted 15,994 investigations. In most ways, the results for the Equivalent Historical Period Comparison paralleled those of the Equivalent Experience Comparison. When the original investigation led to a substantiated finding, a substantiated re-report was 1.35 times more likely to occur within 6 months than when the original investigation was unfounded. Substantiated re-reports were 2.97 times more likely when the caregiver was a parent compared to cases in which a non-parent was the caregiver. Each of the following types of alleged maltreatment was moderately to substantially related to re-reporting substance exposure (1.63), emotional abuse (2.32), lack of supervision (1.97), environmental neglect (1.77) and other neglect (1.50).However, in the Equivalent Historical Period Comparison, sim training was not significantly related to the likelihood of re-report. In other words, when we compared the sim group and non-sim group during the same years in history (even though their experience differed), the sim and non-sim cases were about equally likely to have a re-report.

These results need to be put in context. Investigation is a brief intervention, and no matter how well conducted, it may not be powerful enough to have an impact on re-reports. Bartko and colleagues' research suggests that a substantial trauma-informed intervention may reduce the risk of re-reporting, but this is a fairly substantial intervention specifically targeted to factors that place families at risk for re-reporting.

Chapter 6: Report Conclusion and Recommendations

The COVID-19 pandemic that erupted in the Spring of 2019 persisted throughout the entire FY2020 and restrictions due to the pandemic are still in place now that the fiscal year is ending. The Child Protection Training Laboratory moved entirely to virtual training in August 2020. Most of the data analyzed in this report come from the virtual trainings.

The program evaluation provides data supporting the value of simulation training even when delivered virtually. Trainees' reports from the Daily Experience of Simulation Training (DEST) measure showed increases in confidence in child protection skills during the virtual simulation training weeks. On the post-training survey, a majority of respondents agreed or strongly agreed with the items indicating satisfaction with simulation training. Yet some results suggest that trainees did not respond as positively to the virtual simulation training as previous trainees responded to in-person simulation training. As we discussed in Chapter 3, data from this fiscal year suggest that online training can be a helpful adjunct when in-person training is impossible or impractical, but it is probably not wise to consider it equivalent to in-person simulation training.

Some respondents reported challenging experiences in simulation training. Simulation training is an emotionally powerful experience, and the fact that some people experience difficulties with it should be an expected outcome in both simulation and child welfare work in the field. More research is needed regarding trainees who have challenging simulation training experiences.

Training is designed to increase trainees' learning as a means to influence trainee behavior and thereby produce better results. The analysis of turnover evaluates simulation training in terms of an important trainee behavior: leaving their job as investigators. Turnover is very difficult to

study as an effect of simulation training, because there are so many factors in addition to training that could influence turnover. It cannot be very surprising that we found no relationship between simulation training and turnover given the many different factors that could influence turnover. As desirable as it is to measure investigators' behavior as a way to evaluate simulation training, it is difficult. A promising approach to assessing trainee behavior is to measure the competency of trainees who participate in simulation training. Havig and colleagues developed a reliable measure of rating simulation trainees on three skill dimensions: rapport-building, communication and information-gathering, and safety assessment and ending. CPTA has recently embarked on a plan to adapt this method to assess the competency of its trainees.

It would be valuable to show the relationship between simulation training and client results, because the only purpose ultimately of simulation training is to benefit children and families. However, it is difficult to show this relationship, given the limited degree to which we can measure client outcomes, the variety of factors that can influence client outcomes, and historical trends that could influence any client outcome. Perhaps client's experience of the investigation would be the most important client outcome to measure in relation to simulation training, because the most important lesson of simulation training may be learning to engage clients and maintain the relationship throughout the process of investigation. Of course, even if interviews revealed that families had positive appraisals of sim-trained investigators, without an adequate comparison group, we could not infer that simulation training caused this outcome. There are limitations in using non-sim trained DCFS investigators as a comparison group.

Simulation is a rapidly growing method for training child protective services workers, but it is still early in its development and much remains to be learned through research and program evaluation. Results from the program evaluation of CPTA over a five-year period indicates that simulation training is a powerful and promising method in child welfare that deserves further development and study. The findings of this year's program evaluation suggest that studies examining field outcomes such as turnover and re-report are not likely to be fruitful at this stage, while the immediate impact of simulation training needs to be understood better.

Chapter 1: Introduction

The Child Protection Training Academy (CPTA) is a program to provide experiential learning through simulation training and related methods to professionals from the Illinois Department of Children and Family Services (DCFS). CPTA was developed through a partnership of the University of Illinois at Springfield, the Illinois Department of Children and Family Services, and the University of Illinois at Urbana-Champaign. Detailed information on the development of CPTA and its impact is available in previous program evaluation reports (see below for capsule summaries; links to evaluation reports are in the footnotes). This report provides program evaluation result for CPTA for Fiscal Year 2021.

Adapting to the COVID-19 Pandemic

The COVID-19 pandemic had the country in its grip in FY2021 and the Child Protection Training Academy had to adapt to it. From March to August 2020, simulation training was suspended in both the Springfield and Chicago simulation laboratories because of the risk of Covid-19 infection in any gathering. The CPTA then began virtual simulation training of new child protection investigators in both the Chicago and Springfield laboratories using Zoom. In two simulation sites, the mock family was physically present in the laboratory along with trainers, and the trainees participated remotely via Zoom. This was done for the Knock on the Door simulation in which trainees are making first contact and engaging with the family, and for the Scene Investigation. One of the trainers would wear a head camera and serve as a proxy for the trainee. The proxy trainer would approach the family just as the trainee would have in person, and the head camera would serve as the trainee's "eye", enabling the trainee to observe the home and family. Indeed, the trainee could direct the proxy to aim the camera to obtain whatever view was needed. In this way, the trainee was able to observe safety hazards in the home and conduct a visual inspection of the "baby" (doll) to see if there were signs of physical abuse. Other simulations such as the courtroom simulation were conducted entirely over Zoom.

In this report, we present results from trainings that used virtual methods. We used the same measure (Daily Experience of Simulation Training measure and post-training survey questions) that we used previously to evaluate in-person simulation. Thus we can see if trainees' reactions to the virtual simulations differed from their reaction to the in-person simulations.

Training Supervisors in Problem-Based Learning

The CPTA began to train a new group of DCFS professionals this fiscal year: supervisors of investigators. It conducted several trainings with this group. Rather than using simulations with a mock family , the training team used a number of experiential methods involving exposing supervisors to visual or textual representations of families, and training the supervisors to use problem-based learning to guide their inquiry and decision-making about these families. Problem-based learning is a method in which trainees are presented with problems to solve rather than content to memorize.[1] In problem-based learning, trainees learn to gain the knowledge they need for critical decision making, and avoid relying on unconfirmed hunches

---

[1] Savery, J.R. (2015). Overview of problem-based learning: Definitions and distinctions. In A. Walker, H. Leary, C.E. Hmelo-Silver, & P.E. Ermter (Eds.) *Essential Readings in Problem-Based Learning.* Purdue University Press*.*

and hypotheses that can be grounded in assumptions and biases. The methods for evaluating problem-based learning are in development and we could only gather limited data about the problem-based learning training for supervisors. We anticipate adding a component on the evaluation of this training in the future.

Changes in Investigator Training

The Child Protection Training Academy changed the format of simulation training for investigators during the latter part of this fiscal year. They spread simulation training across parts of two weeks rather than concentrating it one week. This gave CPTA greater flexibility to provide the training during months with holidays that interrupted training. It also gave trainees a weekend break with the idea of reducing fatigue—a particular concern when all the training had to be provided virtually through a computer screen.

CPTA also re-formatted investigator training to include two case scenarios rather than just one. This provided trainees a broader experience and gave them extra practice in family engagement. By requiring different standardized patients to play roles in the two families, this change also reduced each standardized patient's time commitment – an important consideration in a year in which longer time periods interacting with others potentially meant greater exposure to the Corona virus.

An Overview of Previous Program Evaluation Results

The FY2017 evaluation[2] used qualitative methods (observation and interviews) to describe the development of the CPTA and develop a logic model for the program. It also analyzed data from a post-training satisfaction survey (N=154) of program graduates. Respondents were asked a series of questions about whether simulation training had been effective. On every item except "feeling prepared for simulation training," 76% to 84% of respondents strongly agreed. Across seven evaluative questions on simulation training, there were 1,052 positive ratings (99.3%) and only 7 negative ratings (0.7%). Content analysis of open-ended survey items showed that trainees frequently volunteered positive comments on the value of simulation training. Survey respondents recommended extending simulation training to a wider range of topics, professionals, and locations.

The FY2018 evaluation[3] included a qualitative component that examined in greater depth the process of developing the training. Interviews and focus groups with 32 stakeholders pointed to how the abilities of the CPTA team drive simulation training. The simulation trainer had a blend of numerous skills that facilitated simulation training. The standardized patients combined an ability to stay in character and provide feedback with an effective partnership with the

---

[2] Cross, T. P., Tittle, G., & Chiu, Y. (2018). *Program Evaluation of Child Protection Training Academy for New DCFS Investigators: Initial Report*. Urbana, IL: Children and Family Research Center, University of Illinois at Urbana-Champaign.
https://www.cfrc.illinois.edu/pubs/rp_20180131_ProgramEvaluationofChildProtectionTrainingAcademyforNewDCFSInvestigators:InitialReport.pdf
[3] Cross, T. P. & Chiu, Y. (2018). *FY2018 Program Evaluation of Child Protection Training Academy for New DCFS Investigators*. Urbana, IL: Children and Family Research Center, University of Illinois at Urbana-Champaign.
https://www.cfrc.illinois.edu/pubs/rp_20181016_FY2018ProgramEvaluationoftheChildProtectionTrainingAcademyforNewDCFSInvestigators.pdf

simulation trainer. Legal professionals in the courtroom roles were motivated to help DCFS workers improve their skills and emphasized collecting the necessary information, communicating information clearly and accurately, and presenting in a professional manner.

In addition, the FY2018 evaluation surveyed 259 current DCFS investigators; about half of those had received simulation training (sim group) and half had not, because they were hired before simulation training was offered (non-sim group). The sim group reported greater ease in acquiring the skills of evidence-based documentation and testifying in court. Sim-trained investigators also valued the contribution of different simulations to preparing them for their job. The survey also found differences between sim-trained and non-sim trained investigators on their thoughts about leaving their job. Non-sim investigators had four times greater odds of reporting that they were actively looking for a position at another department of DCFS. Non-sim investigators also had more than three times greater odds of reporting that they would leave DCFS as soon as they found another job, once age and experience were statistically controlled.

The FY2019 evaluation[4] included multiple sub studies to examine the implementation and outcomes of simulation training. The CPTA made significant changes to their training model and implemented it on August 20, 2018, and the program evaluation team conducted a qualitative study of the new training model. The evaluation team also implemented a method called the Daily Experience of Simulation Training (DEST) to examine trainees' experience of change over the course of the simulation training week. The analyses indicated that trainees' confidence level for 13 skills significantly increased over the course of simulation training week. Confidence levels were measured on a 7-point scale, with 7 representing maximum confidence. Confidence levels on the last day ranged from an average of 5.7 (work as a DCFS investigator, testify in court) to an average of 5.9 (engage families, assess safety, integrate compassion and investigative skill). Effect size statistics indicate that the increases were large for every confidence item. The program evaluation team also conducted an updated analysis of the post-training satisfaction data. DCFS provided the evaluation team with data from the post training survey between February 2016 and April 2019. Although the ratings of simulation training were consistently positive across the past 4 years, the ratings of simulation training decreased somewhat from FY2016 to in FY2019. On the other hand, the mean satisfaction score for simulation training was higher than the mean for classroom training by one-fifth of a point on the 5-point scale, a difference that was small but statistically significant.

Employee turnover has historically been a problem in child welfare and the quality of training may be one important way of addressing turnover. Using employment data from DCFS Division of Budget and Finance, the evaluation team examined whether DCFS investigators who had received simulation training tend to remain in their jobs longer than DCFS investigators who joined DCFS before simulation training was available and did not receive simulation training. Results using the statistical method of survival analysis indicated that investigators in the non-

---

[4] Chiu, Y. & Cross, T. P. (2019). *FY2019 Program Evaluation of Child Protection Training Academy for New DCFS Investigators*. Urbana, IL: Children and Family Research Center, University of Illinois at Urbana-Champaign. https://www.cfrc.illinois.edu/pubs/rp_20190903_FY2019ProgramEvaluationoftheChildProtectionTraining AcademyforNewDCFSInvestigators.pdf

sim group were significantly more likely to leave their job than those in the sim group in their first two years. At Month 18, 37% of non-sim group had left their job compared to 20% of sim group. At Month 23, the turnover rates for the two groups almost converge. The odds of leaving their job for the non-sim group were 1.8 times greater than the odds of leaving for the sim group, after controlling for other variables. The reduction in turnover during investigators' first two years could reflect the impact of simulation training, The caveat, however, is that the simulation training "era" at DCFS could differ in many ways from the era before simulation training began, so there could be other explanations for differences between non-sim trained investigators (hired before February 2016) and sim-trained investigators (hired after February 2016).

Among the components of the FY2020 evaluation[5] was a qualitative assessment of the implementation of a second simulation laboratory for new investigators that opened in Chicago in April 2019. This component used data from observations of the Chicago laboratory, interviews with key stakeholders of the laboratory and document review. Our report described the history of the implementation in Chicago, and explored how the three key players -- simulation facilitators, the actors, and courtroom professionals -- have implemented simulation training to provide effective learning experiences to trainees. Our comparison suggested that that the Chicago laboratory is a modest re-invention of the Springfield laboratory, using Rogers'[6] terminology on diffusion of innovations. The experience with the Chicago laboratory suggests that expansion can be successful while still needing to deal with challenges to maintaining the capacity and quality of the simulation training program.

As in previous years, the Daily Experience of Simulation Training (DEST) measure in FY2020 shows that the confidence that trainees report increased substantially from the beginning to the end of the simulation training week. The DEST analysis by cohort suggests that the increase in confidence measured by the DEST was very consistent across cohorts. The post-training survey showed considerable trainee satisfaction with simulation training and indicated that many trainees want more time in simulation training. However, the program needs to be aware of the small percentage of trainees who have a negative experience.

Program Evaluation Activities in FY2021

In FY2021, the CFRC evaluation team again used multiple sub studies to assess simulation training. Chapter 2 presents results from FY2021 from the Daily Experience of Simulation Training (DEST) measure. The DEST is an ongoing component of the simulation training program for new investigators and CFRC periodically analyzes DEST data to track changes in trainees' confidence over the course of the one-week training. Analyzing the DEST for different cohorts helps assess whether the effects of simulation training on trainees' confidence is being maintained and is consistent across cohorts. Chapter 3 offers new quantitative results from the post-training satisfaction survey that all new investigators are invited to complete following their Certification Training. The analysis looks at data from February 2020 to February 2021 to assess new investigators' appraisal of the training, and compares these results to those of

[5] Chiu, L., Lee, L. & Cross, T.P. (2020). *FY2020 program evaluation of the Child Protection Training Academy for new DCFS investigators*. Urbana, IL: Children and Family Research Center, University of Illinois at Urbana-Champaign.
[6] Rogers, E. M. (2003). *Diffusion of innovations* (5th Ed.). New York: Free Press.

previous years. Chapter 4 reports results from analyses from a new turnover study that examines whether investigators who have received simulation training are less likely to leave their job than investigators without simulation training.

Chapter 5 presents results from an entirely new analysis. This examines the likelihood of a re-report of maltreatment in CPS investigations. Our hypothesis was that a maltreatment re-report was less likely among cases with investigators who received simulation training, compared to investigators who did not receive simulation training. This follows from the logic model for simulation training[7]. Simulation training is thought to increase investigators' skills, which should lead to better decision-making. Better decision-making by investigators should lead to a greater likelihood that families will receive the intervention they need. This should lead to greater child safety and therefore a smaller chance that the family will be re-reported to DCFS. This is the first analysis in the history of program evaluation of CPTA that links simulation training to family outcomes. Chapter 6 considers all the program evaluation results and discusses the implications for understanding and developing the simulation training program.

---

[7] See Cross, Tittle, & Chiu (2018)

Chapter 2: Daily Experience of Simulation Training (DEST)

Simulation training is thought to increase investigators' preparedness for and confidence in their work, which is in turn thought to have a positive impact on both investigators' experiences and the quality of their work with families. The Daily Experience of Simulation Training (DEST) measure was designed to examine trainees' experience of change over the course of simulation training for their work as investigators in the field. During the week of simulation training, trainees rated their confidence daily on 13 child protection work skills. Findings from the DEST have been reported in the past two evaluation reports for FY2020 and FY2019. This chapter will include results from FY2021.

This is an important time to assess DEST results, because of changes in simulation training during FY2021. As discussed in the Introduction, CPTA implemented a virtual adaptation of its simulation training in August 2020 and conducted several virtual trainings. Most of the DEST results reported here come from these virtual trainings. Moreover, CPTA made other changes to its training for investigators (discussed in more detail in the Introduction). The timeframe of the training week is no longer between Monday and Friday. Instead, it is between Thursday (Day 1) and Wednesday (Day 5). Moreover, instead of using a single case scenario throughout the five days of training, two case scenarios, one for Day 1 and Day 2 training and the other for Day 3 to Day 5 training, were implemented. In this chapter, in addition to updating the results from past reports, we will examine whether there is difference in trainees' confidence level before and after the virtual training as well as before and after the implementation of new training structure.

DEST Revision

We implemented a new version of the DEST in FY2021 (DEST 3.0), in response to the simulation training team's requests (see Appendix A for the DEST 3.0 survey content).[8] We added questions concerning the effectiveness of the debriefing. We also added questions regarding on-the-job training prior to starting simulation training, to see if this may have had an impact on their simulation training outcomes. The suspension of simulation training for five months in 2020 meant that some new trainees began simulation training with substantially greater on-the-job training than previous trainees.

Methods

The DEST includes 13 items measuring trainees' level confidence on different child protection skills. Trainees rate their confidence level on each specific item from 1 (low) to 7 (high). The 13 items are analyzed individually but we also calculate an overall confidence score that is the average of the 13 items. The DEST is implemented at six time points over the course of simulation training week: trainees complete a baseline DEST in the morning of Day 1 and also complete a DEST at the end of each day, Day 1 through Day 5. The baseline DEST also includes questions about trainees' on-the-job-training experience. Between Day 1 and Day 5, the DEST

---

[8] The CPTA director assisted the evaluators with developing the original DEST (DEST 1.0) and a pilot test was conducted in early FY2019. The DEST 1.0 was revised based on the pilot study and the DEST 2.0 version was implemented between December 2018 and March 2020. The DEST was revised again per the training team's request. The current version (DEST 3.0) was launched with the virtual simulation training in August 2020.

also asks trainees to rate the helpfulness of feedback and the effectiveness of the debriefing from the training team, and two open-ended question that asks trainees to share a daily reflection of what they have learned throughout the day.

Every day trainees were given a brief amount of time to complete the DEST, although the DEST was voluntary and trainees were free to decline to participate or terminate participation at any time. Trainers did not know which trainees participated and which did not. The data collected through the secure website were automatically saved on a secure server managed by the Children and Family Research Center

Response Rates

The response rate for the DEST at each time point was calculated by dividing the number of responses (numerator) by the total number of trainees in simulation training (denominator). Between August 24, 2020, and April 28, 2021,[9] a total of 138 trainees participated in the simulation training. All of them filled out the DEST at one time point or more. The DEST data included 737 responses throughout six time points. The daily response rate for the six time points ranged from 73% to 95% across the Chicago and Springfield sites (Table 2.1). As compared to the average response rate of online surveys (34.2%),[10] the weighted average daily response rate of 89% is very high. Out of 138 respondents, 84 (61%) completed the DEST at all six time points. Since a large percentage of trainees completed the DEST, it is reasonable to conclude that results from the DEST measure are representative of trainees, and the measure is being used successfully with investigators receiving simulation training.

Analysis

One-way analysis of variance (ANOVA) was used to compare average confidence scores over time for all respondents, whether or not they had responded at all six time points. Repeated measures ANOVA was used to measure change among those 84 who completed the DEST at each time point. Repeated measures ANOVA is a powerful method for examining change over time of the training week because error variance due to trainee differences is eliminated in the calculation of the F statistic, but it can only be used with trainees who completed the DEST at each time point. The Cronbach's alpha reliability coefficients for the overall confidence score were larger than 0.95 at each of the six time points, which indicates excellent internal consistency among the 13 items in the scale.

---

[9] The training team invited two past participants back for a Beta training of the virtual simulation on the week of August 10 before officially launched the virtual simulation format during the week of August 24, 2020. The two participants also did the DEST for the purpose of testing the new version of the DEST. Therefore, the analysis in the report excluded the data of these two participants.

[10] Poynton, T. A., DeFouw, E. R., & Morizio, L. J. (2019). A Systematic Review of Online Response Rates in Four Counseling Journals. *Journal of Counseling & Development, 97*(1), 33–42. https://doi.org/10.1002/jcad.12233

Table 2.1

DEST Response Rate for Each Time Point

Time Point	All (Respondents=138)		Springfield (Respondents=65)		Chicago (Respondents=73)	
	Responses	%	Responses	%	Responses	%
Baseline	110	80%	57	88%	53	73%
Day 1	118	86%	57	88%	61	84%
Day 2	124	90%	59	91%	65	89%
Day 3	126	91%	63	97%	63	86%
Day 4	129	93%	61	94%	68	93%
Day 5	130	94%	61	94%	69	95%

Results

Changes in Confidence Level

Figure 2.1 shows the changes for the entire sample over six time points for the 13 items of the confidence scale. All 13 confidence items showed a substantial linear increase over the course of simulation week. The average trainee's confidence level increased steadily from baseline to the last day across all 13 items. Confidence levels at baseline (Day 1 morning) ranged from an average of 4.3 (testify in court and work as a DCFS investigator) to an average of 5.2 (engage families and identify family strengths). Confidence levels on Day 5 ranged between an average of 5.8 and an average of 6.1 for those same areas. As Table 2.2 shows, one-way ANOVAs with linear contrasts were statistically significant, indicating a significant linear increase in confidence over the course of the simulation-training week for all 13 skills. Results for the effect size measures eta squared ($\eta 2$) and Cohen's d are presented in Tables 2.2 and Table 2.3. According to Cohen's (1988)[11] guidelines, most of the effect sizes were in the medium to large range (i.e., $\eta 2$ = .05 to .11, or d = .66 to 1.17). Cohen (1992, p. 156)[12] has described a medium effect as "an effect likely to be visible to the naked eye of a careful observer" and a large effect as noticeably larger than a medium effect.

---

[11] Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic.
[12] Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.

Figure 2.1

Confidence Level by Time Point
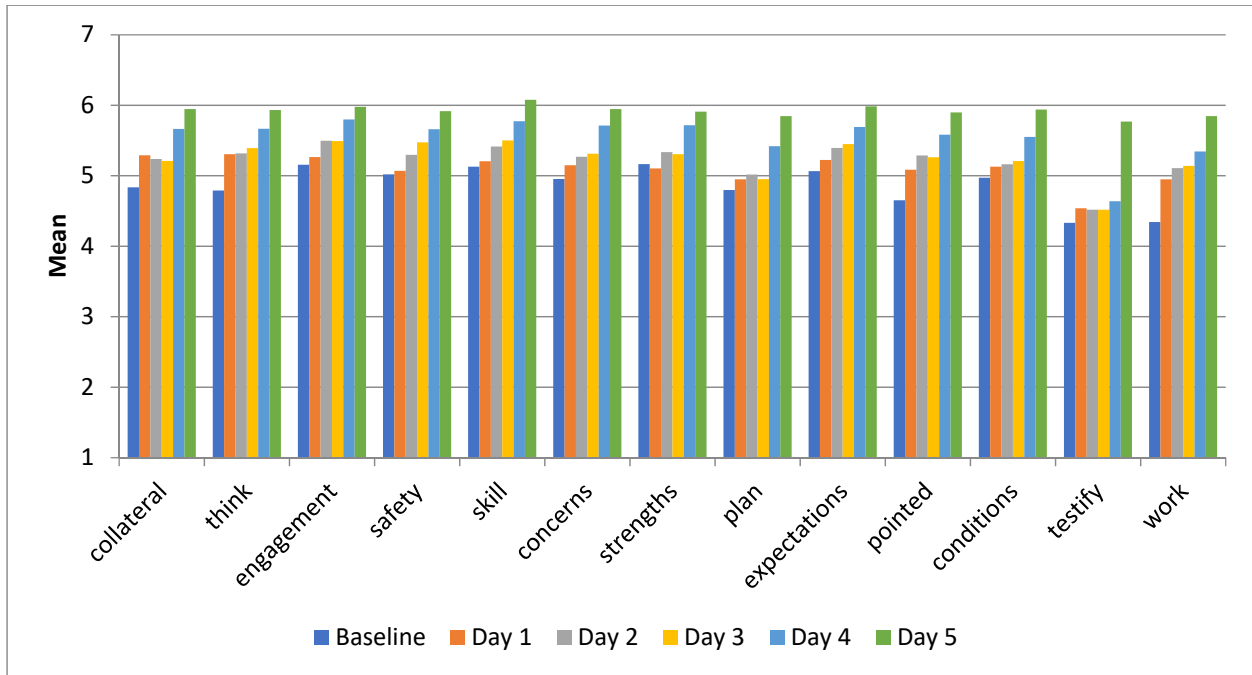
Table 2.2

One-way ANOVA Tests on Confidence over the Course of the Week, Test of Linear Contrasts (all p < .001) (N=732)

Confidence Scale	F (5, 728)	$\eta2$[13]
Gather info from collateral contacts	12.24	0.06
Think critically on facts vs. hypotheses	15.64	0.09
Engage families	9.75	0.06
Assess safety	11.87	0.07
Integrate compassion and investigative skill	13.79	0.08
Address any concerns about family statements and behaviors	14.29	0.08
Identify family strengths	10.40	0.06
Explain need for safety plan and/or protective custody	11.96	0.06
Explain DCFS role and expectations for keeping children safe	10.91	0.07
Answer pointed questions from parents and caregivers	16.74	0.10
Address underlying conditions	11.09	0.06
Testify in court	13.86	0.05
Work as a DCFS investigator	18.87	0.11

---

[13] Computed as $SS_{DEST\ measurements} / (SS_{DEST\ measurements} + SS_{error\ (DEST\ measurements)} + SS_{DEST\ subjects})$. Cohen (1988) has provided benchmarks to define small ($\eta2 = 0.01$), medium ($\eta2 = 0.06$), and large ($\eta2 = 0.14$) effects.

Table 2.3

Statistics for Change between Baseline and Last Day of Simulation Training

Confidence Scale	Baseline		Friday		Cohen's d[14]
	Mean	SD	Mean	SD	
Gather info from collateral contacts	4.83	1.51	5.95	0.98	0.87
Think critically on facts vs. hypotheses	4.79	1.34	5.93	0.93	0.99
Engage families	5.16	1.33	5.98	0.95	0.71
Assess safety	5.02	1.34	5.92	0.94	0.78
Integrate compassion and investigative skill	5.13	1.30	6.08	0.87	0.86
Address any concerns about family statements and behaviors	4.95	1.36	5.95	0.87	0.87
Identify family strengths	5.17	1.26	5.91	0.98	0.66
Explain need for safety plan and/or protective custody	4.80	1.43	5.85	1.02	0.85
Explain DCFS role and expectations for keeping children safe	5.06	1.44	5.98	0.91	0.76
Answer pointed questions from parents and caregivers	4.65	1.38	5.90	0.94	1.06
Address underlying conditions	4.97	1.34	5.94	0.96	0.83
Testify in court	4.33	1.77	5.77	1.08	0.98
Work as a DCFS investigator	4.34	1.51	5.84	1.00	1.17
Total Scale Mean	4.86	1.21	5.92	0.87	1.00

Changes in Confidence Level with the Repeated Measure Sample

A repeated measures ANOVA was conducted with the 84 respondents who completed the DEST at every time point. Differences across time points were statistically significant for all 13 items (Figure 2.2 and Table 2.4). Consistent with the findings in the previous section, the confidence of respondents on performing the 13 investigative skills showed a significant linear increase over the course of simulation training week and the effect sizes were in the medium to large range (i.e., η2 = .05 to .11, or d = .64 to 1.10) (Table 2.4 and 2.5). As Figure 2.2 illustrates, confidence increased steadily for almost all skills across the simulation training week. The skill Testify in Court showed a somewhat different pattern. Average confidence for this skill stayed near baseline until Day 5, when it increased substantially – Day 5 was the day trainees did the courtroom simulation.

---

[14] The rule of thumb on magnitudes of Cohen's is that d = 0.2 are small; 0.5-Medium; and 0.8-Large (Cohen, 1988, 1992).

Figure 2.2

Changes in Confidence Level over 6 Time Points of the Simulation Training Week
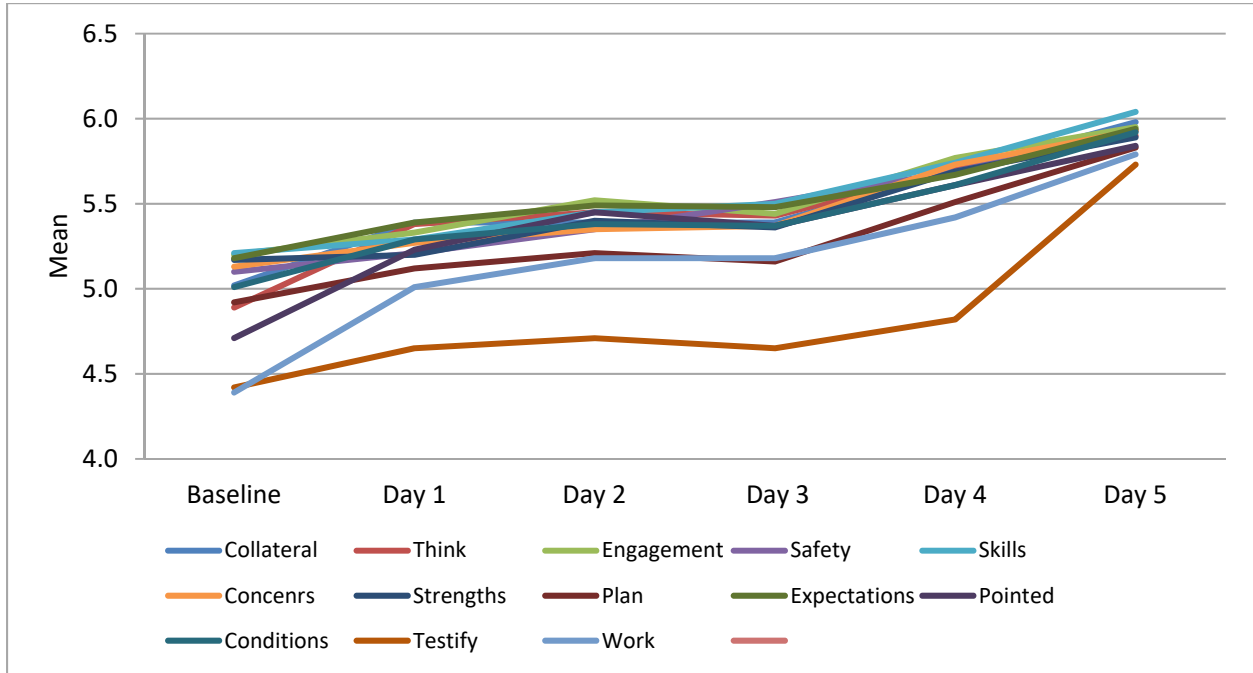
Table 2.4

Repeated Measures Analysis of Variance Test of Linear Effects (all p < .001) (N=84)

Confidence Scale	F	η2
Gather info from collateral contacts	42.50	0.06
Think critically on facts vs. hypotheses	66.58	0.08
Engage families	30.97	0.05
Assess safety	46.03	0.06
Integrate compassion and investigative skill	47.70	0.06
Address any concerns about family statements and behaviors	36.74	0.06
Identify family strengths	32.06	0.05
Explain need for safety plan and/or protective custody	44.53	0.06
Explain DCFS role and expectations for keeping children safe	30.72	0.05
Answer pointed questions from parents and caregivers	57.92	0.10
Address underlying conditions	40.02	0.06
Testify in court	43.94	0.07
Work as a DCFS investigator	71.68	0.11

Table 2.5

Statistics for Changes between Baseline and Last Day of Simulation Training-Repeated Measures Analysis of Variance Sample (N=84)

Confidence Scale	Baseline		Friday		Cohen's d
	Mean	SD	Mean	SD	
Gather info from collateral contacts	5.02	1.34	5.98	0.96	0.82
Think critically on facts vs. hypotheses	4.89	1.27	5.90	0.98	0.89
Engage families	5.20	1.30	5.95	0.98	0.65
Assess safety	5.10	1.25	5.93	0.97	0.74
Integrate compassion and investigative skill	5.21	1.23	6.04	0.95	0.75
Address any concerns about family statements and behaviors	5.13	1.25	5.93	0.95	0.72
Identify family strengths	5.17	1.23	5.89	1.01	0.64
Explain need for safety plan and/or protective custody	4.92	1.30	5.83	1.04	0.77
Explain DCFS role and expectations for keeping children safe	5.18	1.30	5.94	1.01	0.65
Answer pointed questions from parents and caregivers	4.71	1.29	5.84	1.02	0.97
Address underlying conditions	5.01	1.24	5.92	1.00	0.81
Testify in court	4.42	1.63	5.73	1.10	0.94
Work as a DCFS investigator	4.39	1.42	5.79	1.10	1.10
Total Scale Mean	4.95	1.11	5.90	0.95	0.91

Examining DEST Results Across Cohorts

Comparing DEST results across cohorts enables us to see if changes in trainees' confidence have been consistent across trainings. This is a form of quality control. We examined DEST results by training cohort for 18 cohorts from August 2020 to April 2021. Springfield staff trained some cohorts and Chicago staff trained other cohorts. For some cohorts, Springfield and Chicago staff trained separate groups simultaneously – data were pooled across trainings for these cohorts. The sample size of each cohort ranged from 4 to 11. Figure 2.3 depicts the results of the cohorts in order from smallest to greatest change. The blue line shows the mean confidence level (across the 13 skills) at baseline for each cohort and the orange line shows the mean confidence level for each cohort at week's end. Thus, the gap between the blue line and orange line represents the increase in confidence over the course of the week. We can see that there is a noticeable gap for most cohorts between the blue line and the orange line in most cohorts, indicating substantial change in most weeks. Though the sample size of each cohort is small, these results suggests that most cohorts, on average, experienced meaningful increases in confidence during virtual simulation training that resemble the increases respondents in the previous year experienced with in-person training. Four cohorts experienced little or no change; in the three of these cohorts, confidence was already high on average at baseline.

Figure 2.3

Trainee Confidence Levels at the Beginning and End of the Simulation Training Week by Cohort in FY2021
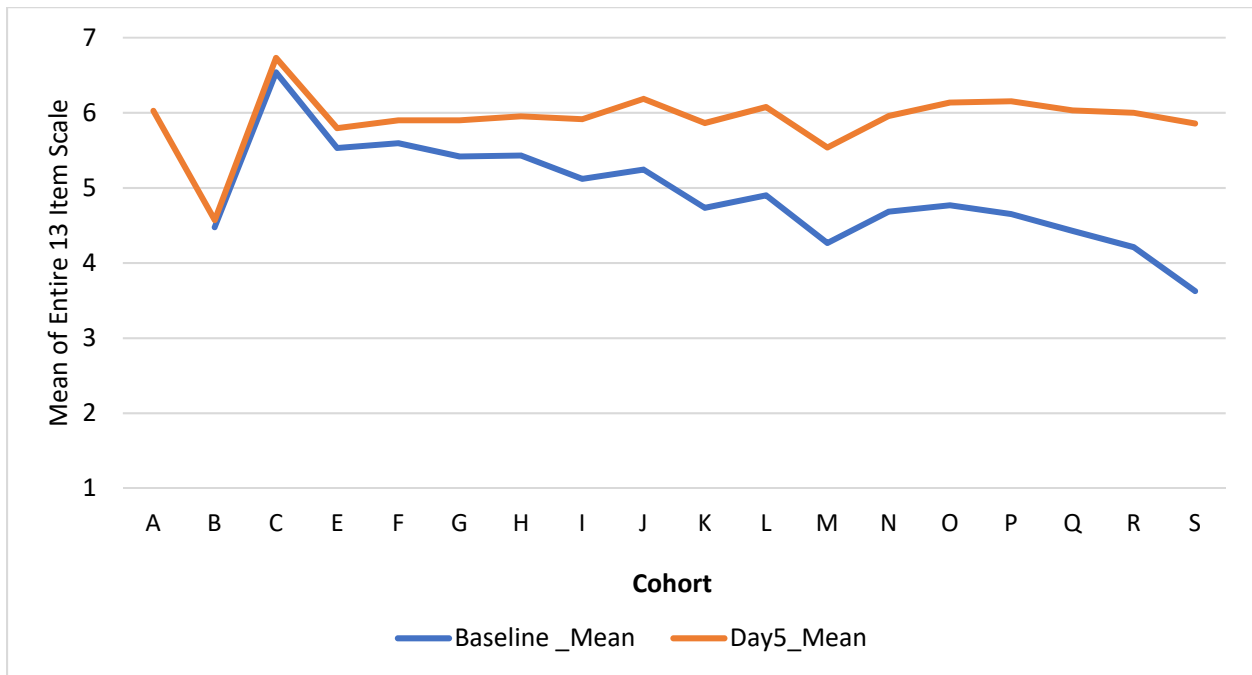
Note: the baseline information for Cohort A was not collected due to a misunderstanding between the evaluators and trainers about the simulation schedule.

On-the-Job-Training and Confidence Level

Due to the pandemic, some trainees might have experienced a greater delay in receiving the simulation training. Consequently, they may have received more on-the-job training (OJT) prior to receiving simulation training than in previous years. This section examines whether the length of time on OJT had an impact on their confidence level during the simulation training week. Most of the respondents (87.1%) were on OJT for four weeks or fewer. Most respondents spent time in their OJT shadowing seasoned investigators, reading related documents, and/or learning about DCFS' Statewide Automated Child Welfare Information System (SACWIS) (Table 2.6).

Analysis with Kendall's tau statistic showed a small but statistically significant relationship between OJT and confidence. Those with longer OJT tended to be slightly more confident than those with shorter OJT, at baseline ($r=.18$, $p<.05$), Day 1 ($r=.19$, $p<.01$), Day 2 ($r=.16$, $p<.05$), and Day 4 ($r=.24$, $p<.01$). Table 2.7 shows Kendall's tau coefficients between OJT and each of the 13 child protection skills, presented separately by day. For most of the skills, except "testify in court," there was a small but statistically significant relationship between OJT and confidence.

Table 2.6

Characteristics of On-the-Job-Training

Time on OJT	n	%	Tasks done during OJT	n	%
None	9	8.3%	Shadowed seasoned investigators	87	79.1%
Less than 1 week	20	18.3%	Read related documents	68	61.8%
1-2 weeks	46	42.2%	Learned about SACWIS	49	44.5%
3-4 weeks	20	18.3%	Worked on investigation reports	17	15.5%
5-6 weeks	3	2.8%	Other	8	7.3%
7-8 weeks	1	0.9%			
More than 8 weeks	10	9.2%			

Table 2.7

Kendall's tau_b: Time on On-Job-Training and Confidence in 13 Skills

	Correlation Coefficient with Time on OJT					
Confidence Scale	Baseline (n=109)	Day1 (n=103)	Day2 (n=102)	Day3 (n=99)	Day4 (n=105)	Day5 (n=101)
Gather info from collateral contacts	.21**	.17*	.23**	.11	.23**	.19*
Think critically on facts vs. hypotheses	.10	.22**	.19*	.20*	.23**	.13
Engage families	.12	.23**	.14	.14	.25**	.18*
Assess safety	.17*	.21*	.19*	.21*	.22**	.17
Integrate compassion and investigative skill	.12	.16	.19*	.12	.23**	.19*
Address any concerns about family statements and behaviors	.15	.18*	.15	.13	.27**	.15
Identify family strengths	.12	.19*	.15	.13	.23**	0.13
Explain need for safety plan and/or protective custody	.24**	.28**	.18*	.18*	.25**	.19*
Explain DCFS role and expectations for keeping children safe	.20*	.27**	.19*	.24**	.24**	.16
Answer pointed questions from parents and caregivers	.19*	.19*	.23**	.18*	.27**	.12
Address underlying conditions	.15	.15	.20*	.09	.26**	.12
Testify in court	.15	.15	.09	.05	.15	.08
Work as a DCFS investigator	.18*	.17*	.15	.15	.19*	.15

*p<.05; **p<.01

New Training Formats and Confidence Level

CPTA changed the training format twice this fiscal year, as mentioned earlier. In response to the Covid-19 pandemic, CPTA delivered the training online. In addition, in February 2021, CPTA

changed the timeframe to Thursday to Wednesday and added an additional mock case. To examine whether these changes influenced trainees' confidence, we conducted one-way ANOVA to test the differences among three timelines: prior to August 2020 (pre-virtual), between August 2020 and February 17, 2021 (virtual), and between February 18 and April 28, 2021 (virtual and new timeframe). We made this comparison for Day 2 and Day 3 confidence levels per the training team's request because Day 2 and Day 3 were the turning points for using the two case scenarios in the new training format. For both Day 2 and Day 3, there was no significant difference in respondents' confidence in the 13 skills among the three timelines.

Appraisal of Feedback and Debriefing

Feedback from the training team during individual and group debriefings is important for facilitating trainees' learning. In the DEST, we asked participants to rate the helpfulness of the training team's feedback and the effectiveness of individual and group debriefings. Based on when trainees received feedback and from whom, we asked respondents to rate the simulation facilitator's feedback every day between Day 1 and Day 4, actors/standardized patients on Day 2 and Day 3, and courtroom professionals on Day 5. Most respondents found the feedback during simulation training either very helpful or helpful (see Figure 2.4). There was no difference in ratings of helpfulness of feedback across roles or time points (Table 2.8). Note that the rating of classroom trainer's feedback was excluded because the classroom trainers did not always attend the simulation training with their trainees after the training format was changed to be virtual.

Figure 2.4

Rating of Trainers' and Actors' Feedback
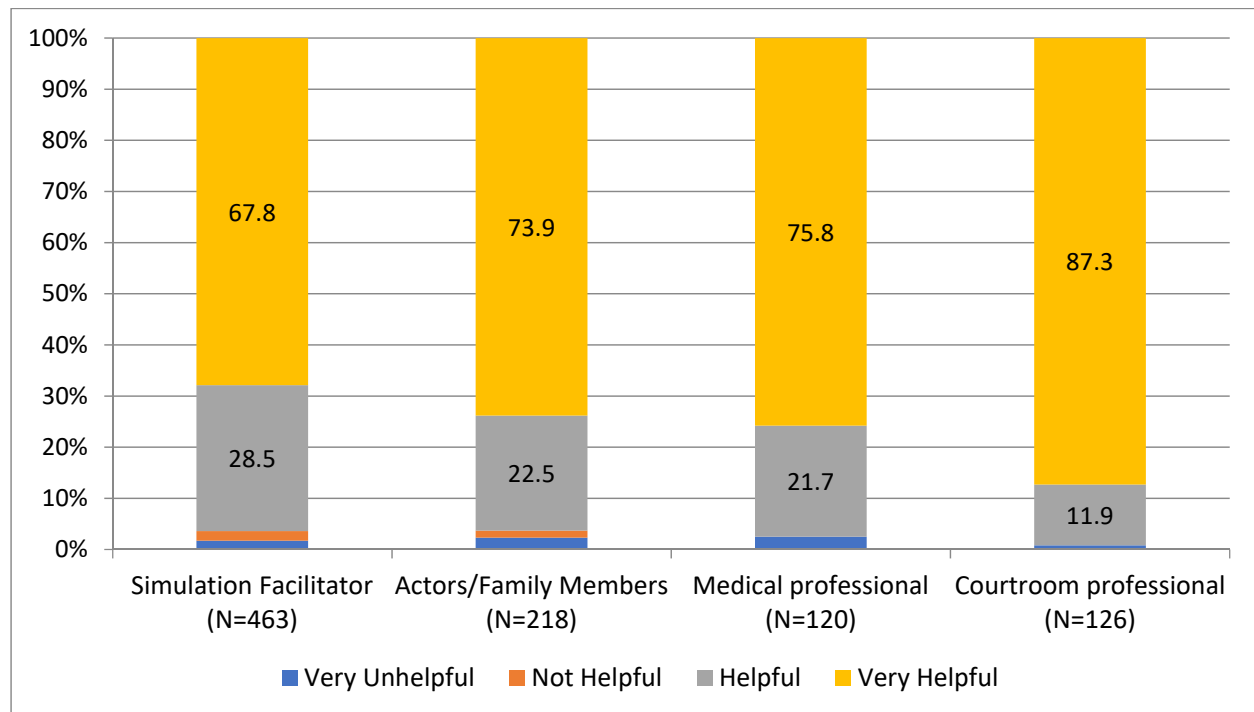
Table 2.8

*Trainees' Ratings of Training Team's Feedback by Days*

	Day 1		Day 2		Day 3		Day 4	
	n	%	n	%	n	%	n	%
Simulation Facilitator								
Very unhelpful	1	0.9%	2	1.7%	3	2.6%	2	1.7%
Not helpful	2	1.8%	0	0.0%	2	1.7%	5	4.2%
Helpful	42	38.2%	27	22.7%	34	29.6%	29	24.4%
Very helpful	65	59.1%	90	75.6%	76	66.1%	83	69.7%
Actors/Family members								
Very unhelpful	-	-	2	1.7%	3	3.1%	-	-
Not helpful	-	-	2	1.7%	1	1.0%	-	-
Helpful	-	-	27	22.3%	22	22.7%	-	-
Very helpful	-	-	90	74.4%	71	73.2%	-	-

DEST 3.0 included new questions about debriefing. Respondents were asked to rate the effectiveness of their group debriefing every day between Day 1 and Day 4 and individual debriefing on Day 2 and Day 3 (individual debriefing was only provided on those two days). Three specific prompts were presented: 1) debriefing identified the areas in which I need to grow; 2) debriefing provoked in-depth discussion that led me to reflect on my skills; and 3) debriefing allowed me to connect with class materials and their practical application.[15] A seven-point rating scale was used, ranging from 1-Extremely ineffective to 7- Extremely effective. Figure 2.9 and 2.10 show that respondents rated the overall effectiveness of debriefing on average between 5.2 and 6.0, indicating both group and individual debriefings effectively facilitated the transfer of learning. Further analyses by training day showed that respondents found the group debriefings on Day 2 (knock on the door simulation), Day 3 (scene investigation simulation), and Day 4 (individual interview and medical simulations) more effective than the group debriefing on Day 1 (calling the reporter simulation) (Table 2.9). In terms of individual debriefing, the average ratings on Day 2 were higher than those on Day 3 (Table 2.10).

---

[15] The Center for Medical Simulation (2009). *Debriefing Assessment for Simulation in Healthcare (DASH).* Authors: Boston, MA. https://www.unmc.edu/academy/community/simulation/wp-content/uploads/sites/5/2017/04/IMSH_2009_DASH.pdf

Figure 2.5

Appraisal of Debriefing Effectiveness
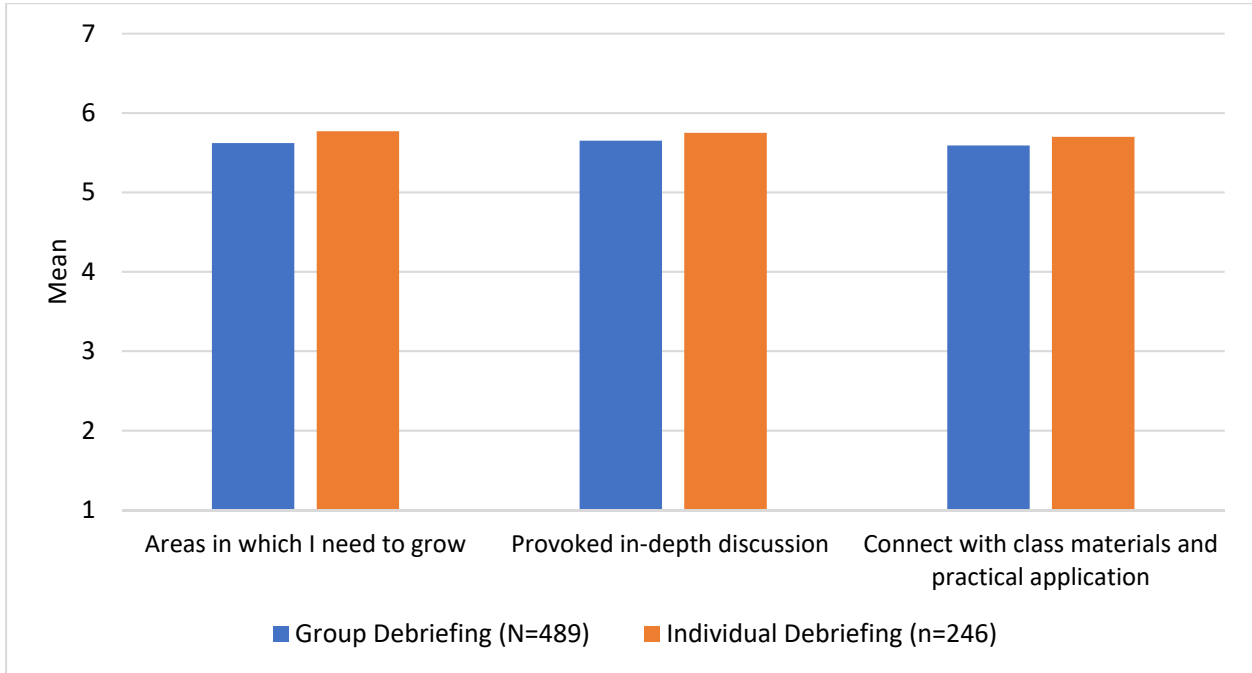
Table 2.9

One-way ANOVA on Rating of Effectiveness of Group Debriefing by Training Day

Group Debriefing	Day 1		Day 2		Day 3		Day 4		*F*	Post Hoc Tests
	M	SD	M	SD	M	SD	M	SD		
Debriefing identified the areas in which I need to grow (n=489)	5.2	1.2	5.9	1.1	5.8	1.1	5.6	1.3	9.14***	Day1 < Day2 Day1 < Day3 Day1 < Day4
Debriefing provoked in-depth discussion that led me to reflect on my skills (n=488)	5.3	1.2	5.9	1.1	5.7	1.2	5.6	1.3	5.88***	Day1 < Day2 Day1 < Day3
Debriefing allowed me to connect with class materials and their practical application(n=486)	5.2	1.2	5.9	1.1	5.6	1.2	5.6	1.3	6.09***	Day1 < Day2 Day1 < Day3

***$p<.001$

Table 2.10

Student's t-test Comparing Ratings on Effectiveness of Individual Debriefing by Training Day

Individual Debriefing	Day 2		Day 3		t
	Mean	SD	Mean	SD	
Debriefing identified the areas in which I need to grow (n=124)	6.0	1.0	5.6	1.3	2.77**
Debriefing provoked in-depth discussion that led me to reflect on my skills (n=123)	6.0	1.1	5.5	1.3	2.74**
Debriefing allowed me to connect with class materials and their practical application(n=121)	5.9	1.1	5.5	1.3	2.35*

*p<.05; **p<.01

Discussion

The Daily Experience of Simulation Training (DEST) provides valuable real-time data on trainees' experience of simulation and is the only evaluation method to date that measures change over the course of the simulation training week. All the simulation training participants completed the DEST at least once during their training week. The high response rates enhance the validity of the results. As in previous years, the DEST in FY2021 continued to show the linear increases in confidence for almost all the 13 skills with the effect sizes in the medium to large range. Confidence in testifying in court increased significantly after trainees had experienced the courtroom simulation. The limitation of the DEST is that it measures trainees' subjective sense of their abilities and is not an objective measure of their skills. Nevertheless, DEST results are important because trainees' appraisal of their skills is likely to have some validity and training is unlikely to be effective if trainees do not believe that their skills are increasing. DEST increases in confidence paralleled those from previous years, suggesting that trainees were able to increase in their confidence in their skills despite the switch to virtual training.

The analysis showed that increases in confidence were mostly consistent across 18 cohorts in FY2021, including both cohorts with Springfield trainees and Chicago trainees. Because sample sizes for this analysis were small and the reliability of individual results is limited, we think it is inadvisable to examine or compare an individual cohort with smaller changes in the DEST. A better use of the cohort results is to conclude that increases in confidence during the simulation training week are typical but not guaranteed, so quality control remains important.

We added several new analyses this year. The results concerning trainees' OJT and confidence level showed that those with longer OJT tended to be slightly more confident than those with shorter OJT. In response to the Covid-19 pandemic, CPTA delivered the training online. In addition, in February 2021, CPTA changed the timeframe to Thursday to Wednesday and added an additional mock case. Our analysis showed on both Day 2 and Day 3 showed that there was no significant difference of respondents' confidence level in the 13 skills resulted from the changes.

In terms of the appraisal of the training team, each member of the simulation team (classroom trainer was excluded in the analysis this year, see above) continued to receive positive feedback from large majorities of trainees this year despite the training being conducted virtually. The results concerning the effectiveness of debriefing with the training team also showed positive results. Trainees rated the overall effectiveness of debriefing on average between 5.2 and 6.0 (on the scale of 1 to 7), indicating both group and individual debriefings effectively facilitated the transfer of learning. The training team could explore reasons for differences in the feedback on debriefing for different days of the training.

Chapter 3: Post-Training Satisfaction Survey

All newly hired child protection investigators participate in Certification Training for Child Protection, which includes five weeks of classroom training and a week of simulation training. DCFS administers an online post-training satisfaction survey on the Certification Training experience to trainees. The previous annual evaluation report analyzed survey responses from February 2019 to February 2020.[16] For this year's evaluation, DCFS provided an updated data set that included survey responses from February 2020 to February 2021. This chapter reports trainees' satisfaction ratings over this time period and also compares results for simulation training, classroom training and for the training program as a whole. It also provides qualitative results from the analysis of open-ended items in the post-training satisfaction survey.

Methods

The post-training survey includes 27 questions about classroom training, 8 questions about simulation training, and 2 questions about the overall training program. Each of the items uses a 5-point Likert scale that ranges from strongly disagree to strongly agree. Descriptive statistics were calculated. Table 3.1 displays the corresponding variable names that the evaluators created for these analyses.

Table 3.1

Simulation training satisfaction questions in the survey

Variable name	Question in the survey
Sim_Prepared	I felt prepared to participate in the SIM lab.
Sim_Environment	The simulation environment was a safe learning environment.
Sim_Learning	I felt the training was conducted in an environment conducive to learning.
Sim_RealisticScenario	The scenario environment was realistic. I was able to incorporate my training into practice.
Sim_RealisticExperience	The SIM lab provided a realistic experience of the challenges I will face when working in the field.
Sim_Confidence	Participating in the scenarios helped to increase my confidence in my role.
Sim_Debriefing	I felt respected during my debriefing.
Sim_Feedback	The debriefing sessions provided valuable feedback.

---

[16] Chiu, Y., Lee, L., & Cross, T. P. (2020). *FY2020 Program Evaluation of the Child Protection Training Academy for New DCFS Investigators.* Urbana, IL: Children and Family Research Center, University of Illinois at Urbana-Champaign; FY2020 report included partial survey data from trainings conducted during February 2020.

Results

Simulation Training Satisfaction

There were 50 survey respondents between February 2020 and February 2021 that were included in the analysis.[17] On a 5-point scale (strongly disagree=1; disagree=2; undecided=3; agree=4; strongly agree=5), the mean of the eight questions ranged from 3.5 (I felt the training was conducted in an environment conducive to learning; the SIM lab provided a realistic experience of the challenges I will face when working in the field; the debriefing sessions provided valuable feedback; and the debriefing sessions provided valuable feedback) to 4.1 (I felt prepared to participate in the SIM lab; see Figure 3.1). Thus for most of these scales, the average score was between "undecided," and "agree." This indicated some degree of satisfaction with simulation training on average. Figure 3.2 shows the distribution of scores for each satisfaction item. On each item, most respondents agreed or strongly agreed, though more than a third of respondents responded "undecided" to "strongly disagree."

There were two training cohorts in which several respondents reported a challenging experience (see below). We compared the mean satisfaction ratings for these two cohorts to the mean satisfaction ratings for all other trainees this fiscal year. Respondents from these two training cohorts had significantly lower mean satisfaction ratings on Sim_Environment through Sim_Feedback (means ranging from 1.38 to 1.88) than the rest of the respondents (means ranging from 3.9 to 4.2).

Figure 3.1

Mean Simulation Training Satisfaction Ratings
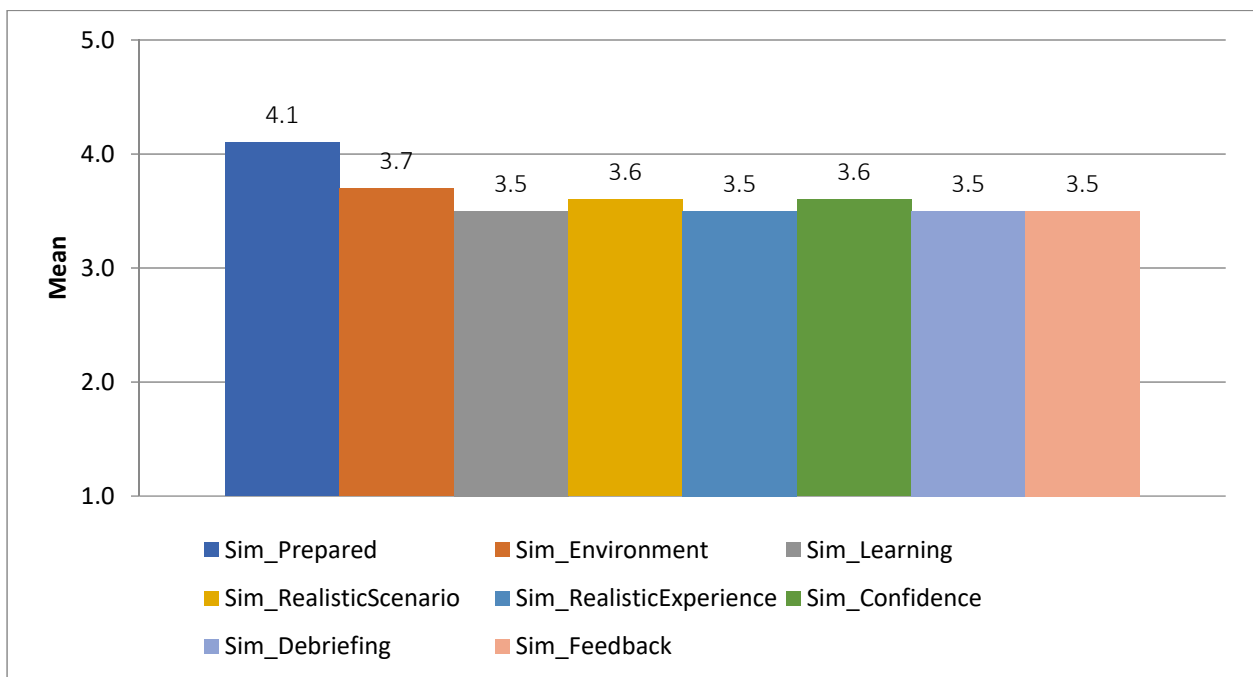
[17] Respondents who indicated that they did not participate in the simulation training were excluded from the analysis.

We compared mean scores on the eight items using a repeated measures analysis of variance, which was not statistically significant (*p*=0.3). Means did not differ across satisfaction items. Again, scores were in the range between undecided and agree on average in this analysis.

Additionally, we compared mean satisfaction ratings for the current fiscal year to mean satisfaction ratings from previous fiscal years. The mean satisfaction ratings for FY2021 were visibly lower than for the previous fiscal years (see Figure 3.3). When we compared the mean satisfaction ratings from FY2021 to FY2020 using an effect size (Cohen's d), the differences were large for Sim_Environment (Cohen's d = 0.8), Sim_Learning (Cohen's d = 1.0), Sim_RealisticScenario (Cohen's d = 0.7), Sim_RealisticExperience (Cohen's d = 0.9), Sim_Confidence (Cohen's d = 0.8), Sim_Debriefing (Cohen's d = 1.0), and Sim_Feedback Cohen's d = 1.0); the difference was small for Sim_Prepared (Cohen's d = 0.1). Overall, the trainees' responses from this year ranged from "undecided" to "agree" on eight scales measuring simulation training; whereas, the responses from the previous years ranged from "agree" to "strongly agree" on the same eight scales.

It should be noted, however, that most of the difference between this year's scores and previous years' scores stems from the very low scores in the two training cohorts in which several people had challenging experiences. If we exclude these two cohorts, mean satisfaction ratings are only slightly lower than in previous years, with only small Cohen's d's.

Figure 3.2

Distribution of Simulation Training Satisfaction Ratings
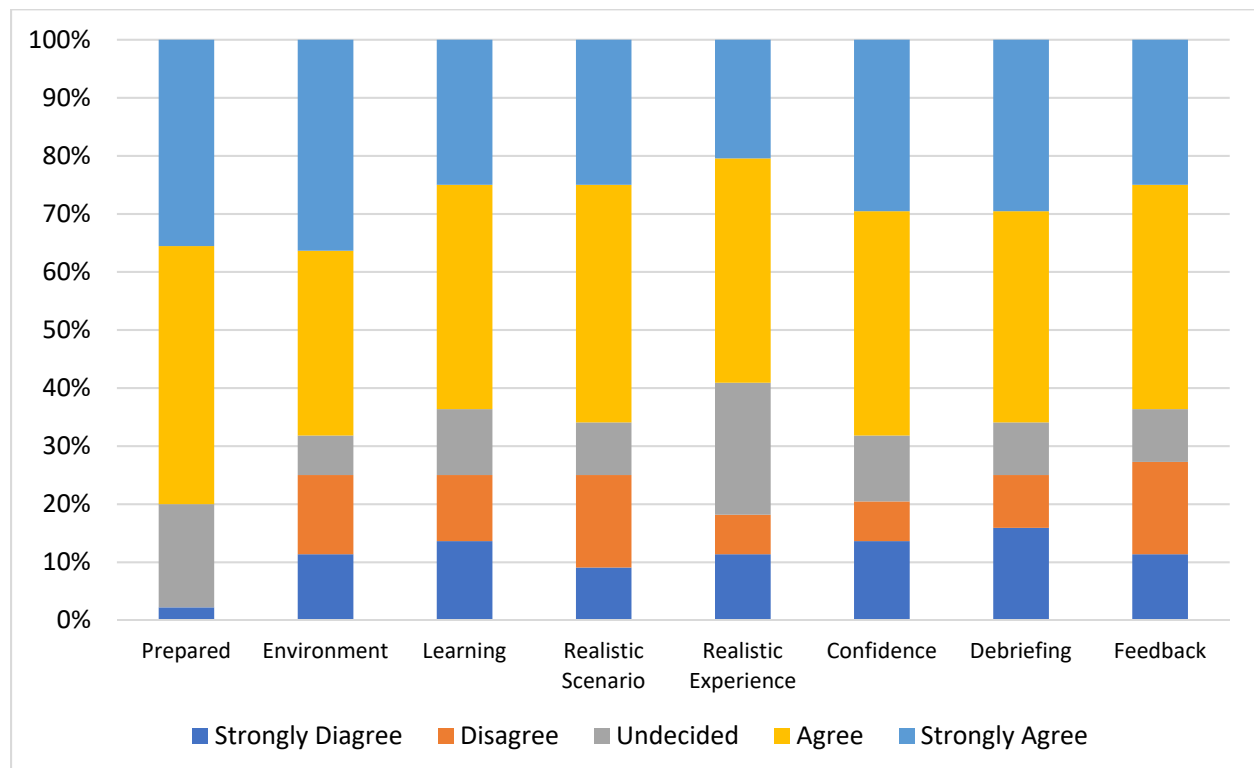
Analysis of Classroom, Simulation, and Overall Program Ratings

Mean satisfaction scores were calculated across the simulation training items, the classroom items, and the overall program items (see Table 3.2). The trainees' satisfaction scores for the classroom and overall program averaged between "agree" and "strongly agree"; while the average satisfaction score for simulation training averaged between "undecided" and "agree" and was significantly lower compared to classroom or overall satisfaction (p=0.01).

Figure 3.3

Simulation Training Satisfaction Rating by Fiscal Year
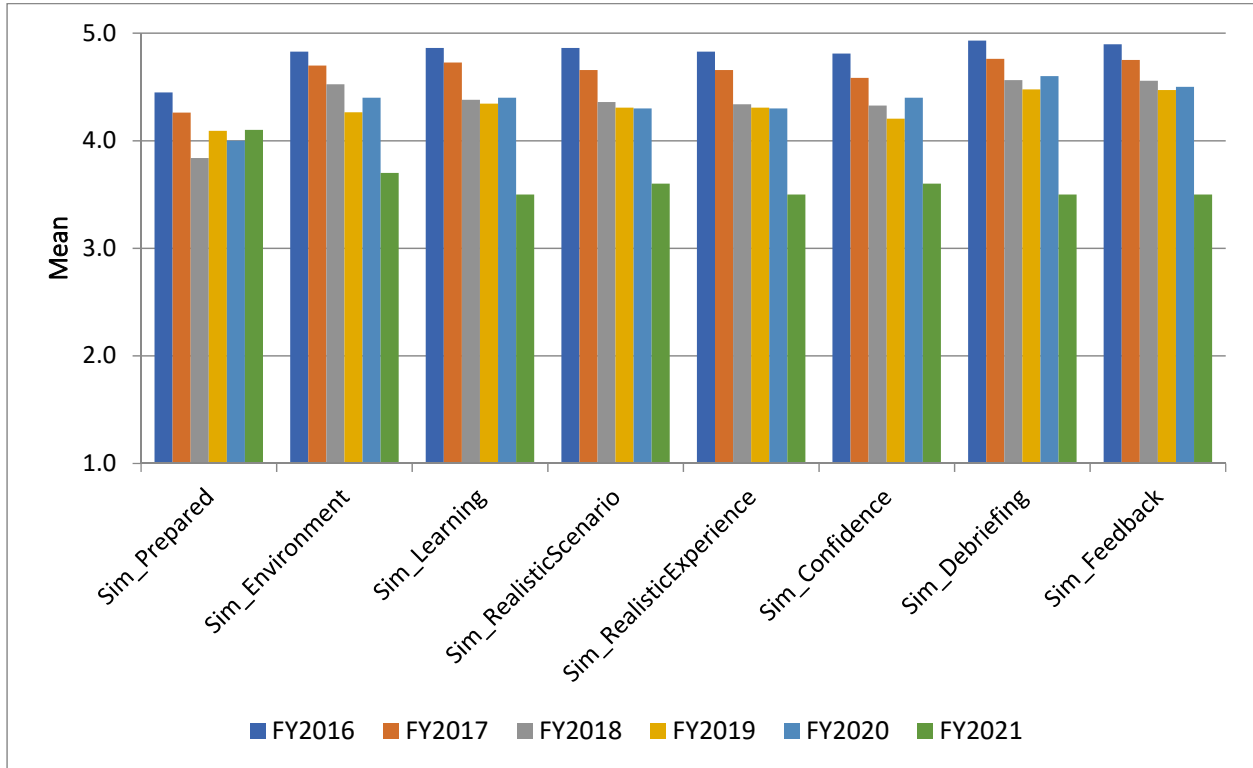
Table 3.2

Descriptive Statistics for Classroom, Simulation, and Overall Program Ratings

Variable name	Total questions in the survey	N	Mean	SD
Classroom Satisfaction	27	50	4.3	0.6
Simulation Satisfaction	8	45	3.6	1.1
Overall Program Satisfaction	2	50	4.4	0.8

Because qualitative data showed that two training cohorts had very dissimilar experiences of simulation training compared to other training cohorts (see below), we compared satisfaction scores by training cohort. The small sample sizes by cohort did not enable us to compute statistical significance tests, but we did find that two cohorts in which several respondents reported a challenging experience had substantially lower scores on the satisfaction items. For example, on the item "The simulation environment was a safe learning environment," five out of eight respondents from these two cohorts recorded "strongly disagree," two recorded "disagree" and one recorded "undecided."

Analysis of Open-Ended Responses

In addition to the Likert-scaled items discussed in the last chapter, the online post-training satisfaction survey includes open-ended items in which trainees can write comments in text. This section presents results of a thematic analysis of text responses to open-ended items on the post-training satisfaction survey data from February 2020 to February 2021.

Two open-ended questions asked about trainees' experience of simulation training: 1) "Comment on this experience" and 2) "Please add a few statements that summarize your experiences in the Simulation Labs to help us improve the scenarios." Because these questions were similar and the answers to the two questions did not differ thematically, we combined respondents' answers to the questions and analyzed the text from both questions together.

We reviewed all the responses and sorted comments based on common themes across multiple respondents. Because respondents were asked the date of their certification training, we were also able to group respondents by training cohort. Altogether 34 trainees provided comments about simulation training. Most responses provided feedback for improving the program; this is consistent with the prompt suggesting the information would be used to improve the simulation scenarios.

Results of Thematic Analysis

Many comments focused on the benefits of simulation training while others described challenging experiences in the simulation laboratory. All of the participant feedback offers opportunities within the existing quality improvement framework to continue building evidence-informed and responsive training for child welfare investigators.

Comments focusing on the benefits were distributed across eight different trainings. Specific benefits mentioned included:

- The importance of application of classroom information to enacted scenarios
- The value of learning about initial engagement with the family
- The opportunity to watch all the debriefing sessions
- The experience of entering the home and seeing the children and home environment
- The realism of the experience (even when done online)
- The opportunity to learn to deal with difficult situations
- The emphasis on the importance of learning one's cases
- The value of receiving a recording of the debriefing to process feedback at a later time
- The importance of a safe learning environment

- The value of helpful and encouraging feedback
- The value of the help, respect, and validation of the simulation trainers.

Comments that described simulation training as challenging came from four different trainings, all done virtually, and most were concentrated in two different trainings (two trainings had one comment each that described simulation training as challenging). Several themes emerged among the respondents who described the training as challenging in terms of their own emotional and cognitive responses. Many of those respondents who experienced simulation training as a challenge commented that they felt disrespected in some way. Many of these respondents also felt that some information they were presented with in simulation training conflicted with their classroom training. The perception was that the instruction in the simulation training was therefore incorrect or inconsistent with what they understand to be policy and procedure, causing confusion and frustration.

Several respondents also reflected on the challenges presented by the actual simulation experiences, pointing to the behavior of one of the actors who they felt was excessively aggressive in a way that interfered with the learning experience. Other comments among those who described a challenging experience expressed concern related to communication, attention, organization, and/or being provided with adequate information. One expressed concern about cultural competence.

Some respondents specifically mentioned the limitations of doing simulation training virtually rather than in-person, though they varied in their appraisal about how much this affected the value of simulation training. A few focused their comment on the limitations of the virtual method and did not report value in the experience. However, most of those respondents who mentioned the limitations of doing the training virtually saw value in the training despite these limitations. Several respondents acknowledged the challenges of doing simulations virtually, but spoke glowingly about the training and felt that the virtual method did not detract from the experience.

Respondents had a number of specific suggestions for improving simulation training, though no one theme predominated. Several respondents expressed the wish to experience simulation training with other child protection tasks or other types of maltreatment, such as sexual abuse. Two respondents felt the training should be focused on one case in the training week.

Discussion

Most of the data from post-training survey Likert-scaled items report satisfaction with simulation training, but, on most items, 30% or more of respondents reported that they were undecided or disagreed or strongly disagreed with the positive statements in the items. Average ratings on these items were noticeably lower than in previous years. However, this is mostly attributable to the particularly low ratings in the two training cohorts with several people with challenging experiences. For the first time, ratings of the simulation training were lower on average than ratings of classroom training and the overall training experience. Again, results from the two cohorts with several people with challenging experiences played a role in this finding.

Many open-ended comments across a number of trainings reported a wide range of benefits of simulation training and communicated a positive experience. There was praise both for the resources provided by the simulation training method and for the personal contributions of the simulation trainers. Comments offering praise outnumbered comments focused on criticism. Comments on benefits represents affirmation of the aims and processes already in place for the simulation training. Supportive feedback is as critical as the comments that indicate areas for continued improvement and help us to recognize what works and what we can continue to build upon.

On the other hand, in a small number of trainings, some trainees had an emotionally challenging experience. This is unusual in the CPTA. Across the five-year history of the CPTA, the vast majority of feedback has been positive, from the post-training survey, from the Daily Experience of Simulation Training (DEST), and from a survey of investigators in the field who had experience simulation training two or three years previously. Indeed, the data from the DEST in Chapter 2 is substantially positive, though a few cohorts showed no evidence of improvement in confidence (in part because the confidence was high for two cohorts to begin with). Nevertheless, it is important to be mindful that simulation training can be emotionally intense, and, as we saw last year, some people have experienced difficulties with it. This should be an expected outcome of both simulation and child welfare work in the field. Simply put, this is hard emotional work. The feeling of being disrespected some respondents reported may reflect the limitations for engagement and rapport-building presented by the new virtual process. The challenges faced by the child welfare workforce and the value of simulated learning are well-established, and these data echo that knowledge.

The perceptions of participants here offer an opportunity for trainers to reflect on the adjustments that may be needed to ensure participants feel engaged with and have a foundational level of trust with them before proceeding into the actual simulation experience The fact that most of the emotionally challenging experiences were concentrated in two trainings implies that the group process in simulation training has not been universally experienced as positive. Taking steps to monitor and support the group process may help forestall challenging training group experiences. We will discuss this further in our concluding chapter below.

We know that child welfare work is incredibly difficult, and simulation can mirror some of those inherent challenges. Additionally, we know that people are often motivated to do child welfare work by their own exposure to child abuse. Studies have revealed that child-serving professionals report higher rates of Adverse Childhood Experiences (ACES) compared to the general population.[18] Child welfare workers also disproportionately report having themselves experienced childhood abuse, a factor that is related to both the experience of vicarious trauma and high turnover among child welfare staff.[19] As such, it is important to continue to keep the

---

[18] Esaki, N., & Larkin, H. (2013). Prevalence of adverse childhood experiences (ACEs) among child service providers. *Families in Society*, *94*(1), 31-37. Lee, K., Pang, Y. C., Lee, J. A. L., & Melby, J. N. (2017). A study of adverse childhood experiences, coping strategies, work stress, and self-care in the child welfare profession. *Human Service Organizations: Management, Leadership & Governance*, *41*(4), 389-402

[19] Festinger, T., & Baker, A. (2010). Prevalence of recalled childhood emotional abuse among child welfare staff and related well-being factors. *Children and Youth Services Review*, *32*(4), 520-526. Nelson-Gardell, D. & Harris, D.

emotional safety and a trauma-informed approach at the forefront of simulation training. The comments by a few trainees around feeling safe are a good reminder that safety is the primary concern in a trauma-informed approach, a framework already endorsed in this model. Trainers can take this opportunity to ensure that the simulation experience is one that includes the understanding that trainees bring their own histories of childhood trauma to the work.

The simulation trainers had to adapt their training to the realities of the COVID-19 pandemic, and were forced to employ virtual methods for conducting simulations rather than forego experiential training altogether. The evidence from the DEST and the post-training survey suggests that virtual trainings still had considerable value for most participants, particularly when we consider trainees who were not in the two training cohorts in which several trainees had challenging experiences. At the same time, post-training survey data also suggest that doing trainings online has some challenges and should not be considered as equivalent to in-person simulation training. It is not known to what extent providing virtual versus in-person training may have impacted participants' experiences. One question to consider is whether some aspect of virtual simulation training may have contributed stress to the group process in the two trainings in which multiple respondents reported challenging experiences, or limited the engagement process. Nevertheless, the data from this fiscal year suggests that online training can be a helpful adjunct when in-person training is impossible or impractical.

One respondent recommended extending simulation training to sexual abuse cases. Perhaps the most challenging area of child welfare work, cases involving the sexual abuse of children can be hard to negotiate on the job and difficult to manage emotionally, with potential impact including vicarious or secondary trauma. Again, also recognizing the likelihood that some trainees themselves are likely adult survivors, this is an area in which added attention may be warranted. One consideration for future training would be to offer the opportunity to simulate a sexual abuse case investigation. Simulation is intended to provide a safe environment in which to learn, make mistakes, practice, and receive feedback without impacting actual children and families. Sexual abuse investigators could benefit greatly from the opportunity to practice this most difficult set of skills prior to doing so in the field.

(2003). Childhood abuse history, secondary traumatic stress, and child welfare workers. *Child Welfare*, *82*, 5-26. Wilke, D. Radey, M., King, E., Spinelli, C., Rakes, S. & Nolan, C. R. (2018) A multi-level Conceptual model to examine child welfare worker turnover and retention decisions, *Journal of Public Child Welfare*, *12*(2), 204-231, DOI: 10.1080/15548732.2017.1373722

Chapter 4: Simulation Training and Investigator Turnover

According to the logic model for the Child Protection Training Laboratory,[20] simulation training is thought to improve child protection professionals' preparation for and confidence in their work, increase job satisfaction and decrease the likelihood that they will leave their job. Staff turnover is a major issue in child welfare: for example, Edwards and Wildeman (2018) found that the median amount of time for a U.S. caseworker between their first and last assigned case was only 1.8 years,[21] and Griffiths and colleagues (2020) estimate that the annual turnover rate in the child welfare workforce range from 15% to 40%.[22] Turnover has numerous negative effects on the quality of child protective services,[23] so any intervention that reduces turnover could improve DCFS services for children and families.

This chapter examines whether simulation training is related to employee turnover. Using two different analytic methods, it asks whether investigators trained using simulation training have stayed in their jobs longer than investigators who were not provided simulation training. On December 18, 2020, the Department of Innovation and Technology at DCFS sent the program evaluators employment data of DCFS investigators from January 1, 2012 to December 31, 2020.

This inquiry updates the CFRC's previous analysis of the relationship between simulation training and turnover. In its FY2019 report, the program evaluation team used the statistical method of survival analysis to compare investigators who had received simulation training (the sim group) to investigators who had *not* received simulation training (the non-sim group) on the likelihood of leaving their job. The investigators without simulation training were significantly more likely to leave their job within two years of their start date than those in the sim group. By Month 18 following their start date, 37% of the non-sim group had left their job compared to 20% of sim group. The odds of leaving their job for the non-sim group were 1.8 times greater than the odds of leaving for the sim group, after controlling for other variables. By Month 23, however, the rate of leaving their job almost converged for the sim and non-sim groups. These results suggested that sim-trained investigators were less likely than non-sim-trained investigators to leave their job within two years of starting it.

---

[20] Cross, T.P., Tittle, G., & Chiu, Y. (2017). *Program evaluation of simulation training for new DCFS investigators: Initial report*. Children and Family Research Center, University of Illinois at Urbana-Champaign. https://cfrc.illinois.edu/pubs/rp_20180131_ProgramEvaluationofChildProtectionTrainingAcademyforNewDCFSInvestigators:InitialReport.pdf

[21] Edwards, F., & Wildeman, C. (2018). Characteristics of the front-line child welfare workforce. *Children and Youth Services Review*, *89*, 13–26. https://doi.org/10.1016/j.childyouth.2018.04.013

[22] Griffiths, A., Collins-Camargo, C., Horace, A., Gabbard, J., & Royse, D. (2020). A new perspective: Administrator recommendations for reducing child welfare turnover. *Human Service Organizations: Management, Leadership & Governance.* Advanced online publication. https://doi.org/10.1080/23303131.2020.1786760

[23] Lee, L., Tran, S., Braun, M.T., LaSota, R., & Fuller, T.L. (2021) *Child Welfare Workforce Task Force: Literature review, employer survey, and recommendations.* Children and Family Research Center, University of Illinois at Urbana-Champaign. https://cfrc.illinois.edu/pubs/rp_20210205_ChildWelfareWorkforceTaskForce:LiteratureReviewEmployerSurveyandRecommendations.pdf

The program evaluation team had a related result in a survey of investigators conducted in FY2018. The odds that a non-sim investigator checked "yes" on the question "I am actively looking for a position at another department of DCFS" were 4.2 times greater than the odds that a sim investigator checked "yes" on that question (p < .05). The odds that a non-sim investigator checked "yes" on the question "As soon as I can find a better job, I will leave DCFS" were 3.5 times the odds that a sim investigator did so (p = .06).

Interpreting the differences between simulation-trained and non-simulation trained investigator requires a significant caveat, however. Simulation training for all new DCFS investigators began in February 2016. No investigator hired before 2016 received simulation training, whereas all new DCFS child protection investigators hired February 2016 or later received simulation training. Thus receiving simulation training is confounded by history: there is a non-simulation training era and a simulation training era. Differences between sim and non-sim investigators could be due to simulation training, but they could also be due to any other differences between the two eras, such as differences in caseload, working conditions, the characteristics of the new investigators who were hired, and so forth. **Thus comparing sim and non-sim investigators involves considerable ambiguity that no statistical procedure can resolve.** The analysis this fiscal year again compared sim investigators to non-sim investigators, and again this caveat applies to our current results.

In the current analysis, we again compared investigators who started to work at Division of Child Protection (DCP) after February 2016 (this group received simulation training) to investigators who started to work at DCP before February 2016 (this group did not receive simulation training). We obtained an updated data file with employment data from DCFS. This file differed from the parallel data file we received from DCFS in FY2019. See Table 4.1 for information on the two data files. The new data set included more recent data, of course, but it was also a more complete data set for earlier periods as well. **Again the caveat applies that a comparison of sim and non-sim investigators is confounded with era and we cannot resolve the ambiguity this introduces.** The updated analyses in this report explore this question using more up-to-date and complete data and utilize new statistical methods.

Table 4.1

Number of Investigators by Year in the Two Data Files

Investigator DCP start year	Data received in FY2019	Data received in FY2021
2012	-	130
2013	-	366
2014	59 (from February)	123
2015	29	66
2016	141	173
2017	154	174
2018	21 (January)	180
2019	-	264
2020	-	27 (until February)
Total Sample	**404**	**1,503**

Interrupted Time Series Analysis

ITS Sample

One method we used to compare sim-trained and non-sim-trained was interrupted time-series analysis (ITS). This involves looking at the trend in turnover rates over time for investigators and seeing if the introduction of simulation training changed ("interrupted") the trend in turnover rates. One advantage of ITS is that we can see overall trends over time in turnover in addition to the comparison between investigators trained in the sim training era and those trained before that. For this analysis, we defined turnover as leaving an investigator position in the Division of Child Protection (DCP) within 12 months of starting it.

In order to clean the data and to select only valid cases, the following steps were taken:

1. We excluded cases in which DCFS left date was recorded as earlier than DCP start date.
2. We also excluded cases in which the difference between DCP start date and DCP left date was less than or equal to 30 days.
3. We deleted cases for investigators hired between January 2016 and February 2016 because we could not determine with precision whether or not they received simulation training.
4. We also needed to aggregate multiple data entries from the same investigator. Some investigators had more than one record in the data file because they had multiple positions with DCFS (e.g., got promoted from a child protection specialist to a supervisor position during the study observation period). The related date variables associated with the change of the position were aggregated so the final sample only included unduplicated investigator records.

All the newly hired investigators after February 2016 have been required to participate in simulation training. Therefore, the *sim group* included all DCFS investigators who started to work at DCP between March 2016 and February 2020. The *non-sim group* included DCFS investigators who started to work at DCP between January 2012 and December 2015 as simulation training had not yet been implemented, and they did not receive it. The total sample consists of 1,412 investigators (669 in the non-sim group and 743 in the sim group). The majority is female (78.8%), Black (44.1%) or White (42.8), with a degree other than social work (73.7%), had a bachelor's degree (80.0%), and an average age of 48.6 years (SD=10.85). Around 19.1% of investigators left their position within the 12-month observation period (Table 4.2).

Table 4.2

Sample Description (N=1,412)

Variable	N	%
Gender		
Female	1,112	78.8%
Male	300	21.2%
Race/Ethnicity		
Black	622	44.1%
White	605	42.8%

Hispanic	150	10.6%
Other	35	2.5%
Social Work Degree		
No	1,040	73.7%
Yes	372	26.3%
Education Level		
High School or Some College	68	4.8%
Bachelor's degree	1,129	80.0%
Master's	210	14.9%
PHD/MD/Doctoral	4	0.3%
Other	1	<0.01%
Position		
Child Protection/Welfare Specialist	1,150	81.4%
Public Service Admin (Supervisor)	235	16.6%
Children and Family Services Intern	27	1.9%
Cohort		
Sim	743	52.6%
Non-sim	669	47.4%
Turnover		
No	1,143	80.9%
Yes	269	19.1%

ITS Data Analysis

Each ITS analysis occurred in a series of steps:

1. ITS analysis requires repeated measures over time (i.e., the time series); the current study used quarterly time points (i.e., every three months), resulting in 33 quarters between Q1 of 2012 to Q1 of 2020. The first step was to count the number of investigators hired during each quarter.
2. We calculated the number of investigators who experienced turnover in each quarterly time point during the 12-month observation period. We determined that turnover occurred when the variable "Left DCP Date" was coded as "yes" and was within 12 months of their "DCP Start Date."
3. For each quarter, the turnover rate was calculated by dividing the number of investigators who started in that quarter and experienced turnover within 12 months by the total number of investigators hired in the respective quarter.
4. We then used the results of the analyses to estimate the effect of several variables on the outcome variable:
   a. The intercept of the time series before the intervention (simulation training); in other words, the baseline level of the outcome of turnover
   b. The non-sim-training trend (a mathematical slope)
   c. The change in intercept (or level) after the introduction of sim training
   d. The change in slope or trend after the introduction of sim training

e. The p-values associated with the parameter estimates indicate whether the relationship of a variable to the outcome was statistically significant.
5. We also conducted tests to make sure the data met relevant mathematical assumptions for regression analysis.
6. Interrupted time series analysis assumes that error terms associated with each observation are uncorrelated: the independence assumption. The rates in quarters that are proximal to one another may be more similar than quarters that are further away in time, which would be an indication of autocorrelation. We tested the independence assumption with the Durbin-Watson statistic, which tests whether the turnover rate in one quarter can be predicted by the previous quarter.
7. We tested for autoregressive conditional heteroscedasticity effects by examining changes in variance across time by using lag windows ranging from 1 to 12. This tests for the assumption of equal variability across time.
8. We also prepared plots of the quarter-by-quarter changes in outcomes. These plots include the fitted regression lines that show the observed trends and lines that estimate what rates would have been if the intervention had not been implemented.
9. A limitation needs to be acknowledged: To make the data as timely as possible, we included in our analysis investigators who started in Quarter 1 of Calendar Year 2020 even though our data did not allow us to have a full 12-month observation period for them.

ITS Results

Effect of Simulation Training on turnover rate of investigators within one year of employment

There were 1,412 investigators in Illinois who started employment between Q1 2012 to Q1 2020, and 269 of them left their job within one year of employment. Table 4.3 shows the number of investigators that were hired in each quarter and the number of those who left the job. Simulation training started in the first quarter of 2016, which is highlighted in yellow.

Data were first examined to check whether assumptions were met. The Durbin-Watson statistic was not statistically significant, indicating that the assumption of no significant autocorrelation was met (Durbin-Watson statistic = 2.15, p = 0.55). Wong and Li's test showed no significant autoregressive conditional heteroscedasticity effects.

Table 4.3

Turnover rate of investigators within one year of employment in Illinois (Quarter 1, 2012 – Quarter 1, 2020)

Year	Quarter	# of investigators	# of investigator who turnover within 12 months	Turnover Rate
			Non-Sim Group	
2012	Q1	40	4	10%
2012	Q2	49	5	10%
2012	Q3	4	0	0%
2012	Q4	30	1	3%
2013	Q1	278	35	13%
2013	Q2	48	8	17%
2013	Q3	16	1	6%
2013	Q4	17	0	0%
2014	Q1	41	5	12%
2014	Q2	38	11	29%
2014	Q3	20	1	5%
2014	Q4	23	7	30%
2015	Q1	17	4	24%
2015	Q2	14	4	29%
2015	Q3	16	4	25%
2015	Q4	18	2	11%
			Sim Group	
2016	Q1	6	0	0%
2016	Q2	52	2	4%
2016	Q3	33	5	15%
2016	Q4	20	2	10%
2017	Q1	42	4	10%
2017	Q2	25	4	16%
2017	Q3	53	7	13%
2017	Q4	52	8	15%
2018	Q1	41	7	17%
2018	Q2	44	9	20%
2018	Q3	42	5	12%
2018	Q4	49	12	24%
2019	Q1	41	16	39%
2019	Q2	46	13	28%
2019	Q3	125	54	43%
2019	Q4	45	17	38%
2020	Q1	27	12	44%

The results of the interrupted time series regression analysis, shown in Table 4.4, indicate that prior to simulation training (the intercept), turnover rates within one year of employment were about 3.3%, and there was a significant trend or quarter-to-quarter change in turnover rates prior to the introduction of the intervention (t = 3.12, p = 0.004 for baseline trend). Following the implementation of the simulation training, the turnover rate decreased significantly by 24.8% (t = -4.82, p < 0.001), and the post-interruption trend in referral rates increased significantly over time (t = 2.20, p = 0.036). In other words, there was a significant drop in the one-year turnover rate following the introduction of simulation training followed by an increase over time.

The results are graphically displayed in Figure 4.1. The blue line shows the observed turnover rates, the grey line represents the fitted regression line estimating what turnover rates would have looked like without the introduction of simulation training, and the orange line represents the estimated turnover rates with simulation training. Figure 4.1 shows the significant decrease in turnover rates following the introduction of simulation training in January 2016 as well as the increase from January 2016 through the first quarter of 2020.

Table 4.4

Parameter estimates predicting turnover rate of investigators in Illinois

Variable	Parameter Estimate	Standard Error	t Value	p
Intercept	0.0334	0.0388	0.86	.396
Baseline trend/slope	0.0125	0.0040	3.12	.004
Level change	-0.2484	0.0515	-4.82	<.0001
Trend change/slope	0.0119	0.0054	2.20	.036

**Discussion of the ITS Analysis**

The results in the current ITS analysis resemble in some ways the survival analysis conducted in 2018 (see above). DCFS investigators who received simulation training were less likely to leave their job within a year than DCFS investigators who had not received simulation. To repeat, again the caveat applies that a comparison of sim-trained and non-sim-trained investigators is confounded with era and we cannot resolve the ambiguity this introduces.

This current ITS study found that turnover increased for those investigators starting in 2014, went down for those investigators starting in 2016, and went back up again for those investigators starting in 2019. There could be historical reasons for these changes – for example, perhaps COVID influenced the increase in turnover for those investigators starting in 2019.

The turnover rate for those who had simulation training (orange line) is less than for those who did not have simulation training (grey line) for the period of time studied. However, there is no way to be sure whether this difference is due to simulation training versus other historical events that were contemporaneous with the introduction of simulation training. Ultimately there is no way to distinguish between effects due to simulation training versus effects due to other changes during the sim training era.

Figure 4.1

Observed and estimated turnover rates of investigators within one year of employment with and without simulation training in Illinois
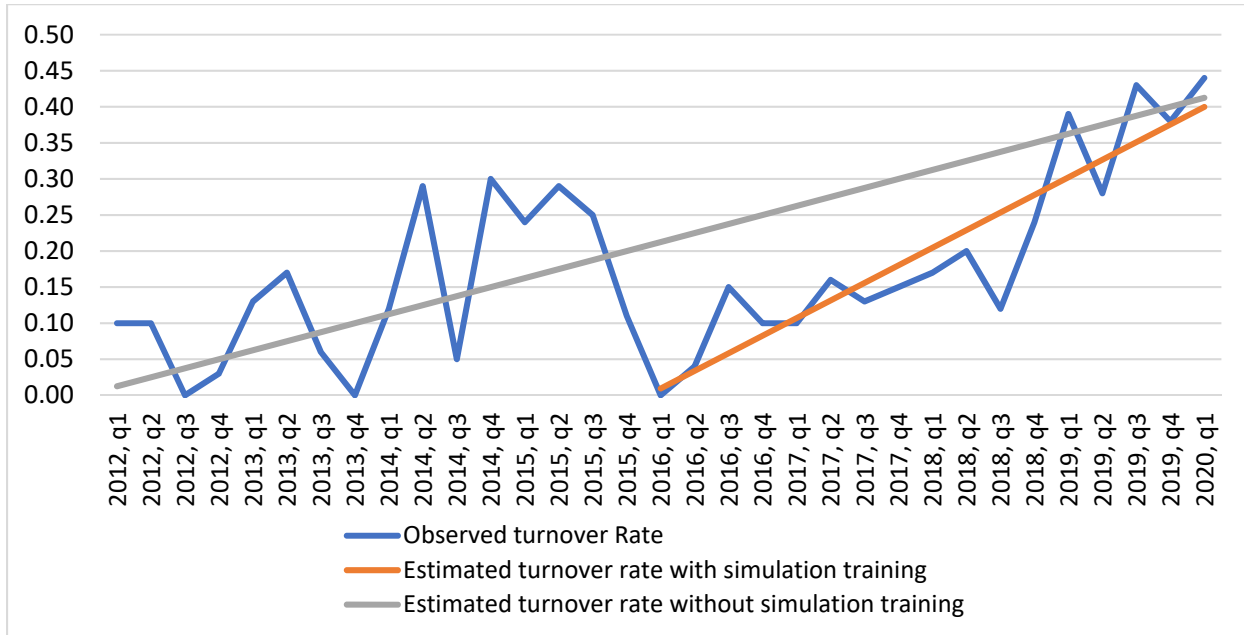
Survival Analysis of Matched Samples

As we did in FY2019, we also conducted a survival analysis to compare the sim group and non-sim group on turnover. Survival analysis yields the probabilities that staff will leave the investigator positions during a given time interval. From survival analysis, we developed a "life table" that shows the probability of "survival" over time, that is, the probability that investigators will remain in their job.

Propensity Score Matching

We added an additional feature to the current survival analysis that we did not use in FY2019. To control for differences between sim and non-sim investigators on such variables as experience and age, we created samples in which each sim investigator was matched with a comparable non-sim investigator. This created matched samples that were similar on almost every variable. We used a method called *propensity score matching (PSM)*. We used a logistic regression model including multiple variables to calculate a propensity score, representing an estimate of the likelihood that an investigator would be a sim-trained investigator.[24] Sim and non-sim investigators with similar propensity scores were matched. We used nearest neighbor matching within caliper and without replacement as a matching algorithm. Propensity score matching is a more thorough way of controlling for confounding factors differentiating two groups than the use of covariates that we employed in FY2019.

---

[24] Rosenbaum, P.R., & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Table 4.5 shows the characteristics of the non-sim and sim groups prior to matching and after the matched samples were created. Note the differences in the pre-match sample between the two groups on whether they had a social work degree, their level of education, and their years of experience. In contrast, the two groups in the post-match sample are similar on most variables except for race – the sim group in the matched sample has a higher proportion of investigators who are White.

Table 4.5

Characteristics of the Pre-Match and Post-Match Samples

	Non-sim	**Sim**	**Difference**
Pre-match sample (N=1,340)[25]	(n=638)	(n=702)	
Age at DCP start	46.2	41.4	4.8
Gender			
Male	22.4%	19.8%	3%
Female	77.6%	80.2%	-3%
Race			
White	41.4%	45.4%	-4%
Black	46.2%	42.0%	4%
Hispanics	10.3%	10.0%	0%
Other	2.0%	2.6%	-1%
Social Work Degree	35.9%	17.9%	18%
Education level			
Higher school or associate degree	0.6%	8.4%	-8%
Bachelor's degree	78.2%	81.6%	-3%
Master's degree or higher	21.2%	10.0%	11%
Position			0%
Children and family service intern	3.0%	5.7%	-3%
Child protection specialist	78.1%	88.3%	-10%
Supervisor or higher	19.0%	6.0%	13%
Experience in DCFS prior to DCP (days)	4,057	1,307	2,750
Post-match sample (N=808)	(n=404)	(n=404)	
Age at DCP start	43.77	41.25	2.5
Gender			
Male	21.0%	23.3%	-2%
Female	79.0%	78.7%	0%
Race			
White	47.0%	62.1%	-15%
Black	39.6%	26.5%	13%
Hispanics	10.9%	7.4%	4%
Other	2.5%	4.0%	-2%

---

[25] The ITS and survival analysis have different samples sizes because the survival analysis sample excluded the investigators who started their DCP position in 2020.

Social Work Degree	34.9%	31.2%	4%
Education level			
Higher school or associate degree	1.0%	1.0%	0%
Bachelor's degree	77.7%	83.2%	-5%
Master's degree or higher	21.3%	15.8%	6%
Position			
Children and family service intern	4.7%	5.7%	-1%
Child protection specialist	80.9%	83.9%	-3%
Supervisor or higher	14.4%	10.4%	4%
Experience in DCFS prior to DCP (days)	2,707	2,186	521

Survival Analysis Results

We conducted the survival analysis with the propensity score matched samples, while also statistically controlling for the effects of the following variables age at starting in the DCP, gender, race, social work degree, education level, current position, and experience in DCFS prior to working in DCP. Figure 4.2 shows the life table from the survival analysis. Through the first 6 months on the job, sim-trained investigators were somewhat less likely to leave their job. From the ninth to the twelfth month on the job, however, sim trained investigators were more likely to leave their job. It should be noted that the first 12 months for the sim investigators and the first 12 months for the non-sim investigators occurred in different historical eras. For the non-sim investigators, the first 12 months occurred in the years 2012 to 2016. For the sim investigator, the first 12 months occurred in the years 2016 to 2020. Note that the time period for the sim investigators includes time in which the COVID-19 pandemic affected the country.

Figure 4.2

Life Table of Survival Analysis: Turnover Within 12 Months Between Non-simulation and Simulation Groups (N=808)
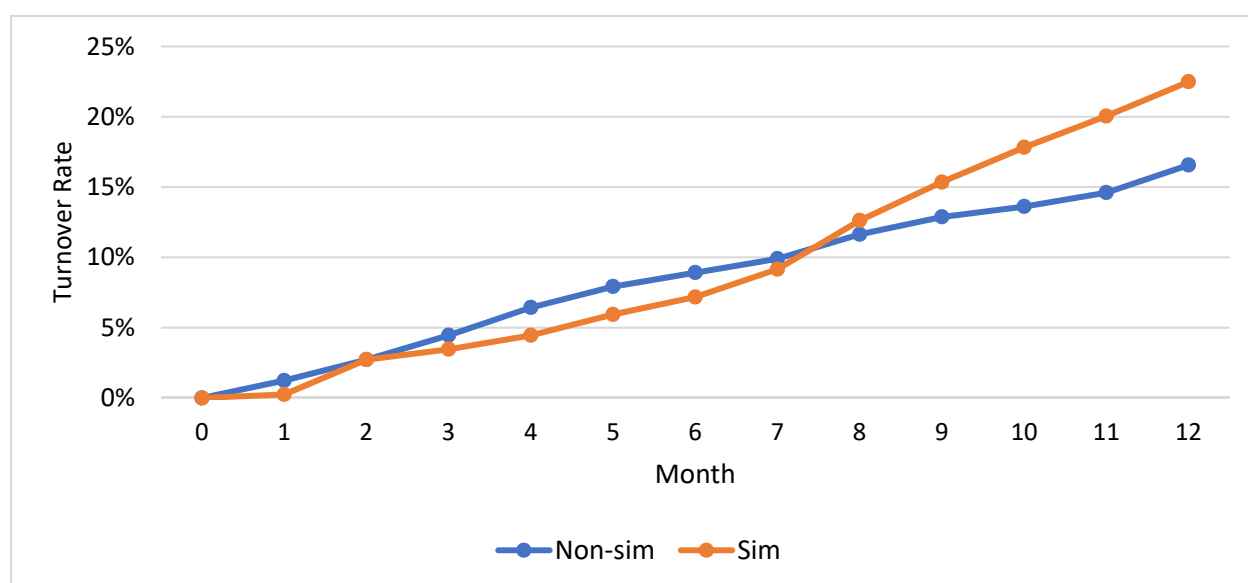
Table 4.6 shows the results from the Cox regression equation calculated as part of the survival analysis. This shows that the odds of the sim group leaving their job within 12 months were 1.50 greater than the odds of the non-sim group doing so. Other significant effects may be of interest too. The odds that female investigators would leave within 12 months were 0.68 the odds of male investigators doing so. The odds that Hispanic investigators would leave were 1.99 times the odds that White investigators would do so. The odds that a child protection specialist would leave were 0.52 the odds that a child and family services intern would leave.

Table 4.6

Cox Regression: Predicting Turnover Rate within 12 months (N=808)

|	Coefficient	SE	Exp(b)
Age at DCP start	0.013	0.010	1.014
Gender			
Male	-	-	-
Female	-0.393	0.188	0.675*
Race			
White	-	-	-
Black	0.233	0.191	1.263
Hispanic	0.686	0.257	1.986**
Other	0.431	0.407	1.539
Social Work Degree	-0.409	0.203	0.664*
Education level			
Higher school or associate degree	-	-	-
Bachelor's degree	-0.866	0.596	0.421
Master's degree or higher	-0.773	0.635	0.462
Position			
Children and family service intern	-	-	-
Child protection specialist	-0.641	0.323	0.527*
Supervisor or higher	-0.331	0.415	0.719
Experience in DCFS prior to DCP (days)	0.000	0.000	1.000
Cohort			
Non-sim	-	-	-
Sim	0.418	0.171	1.519*
−2 log Likelihood	1932.746		
X²	35.392		
df	12		
Sig.	<.001		

*p<.05; **p<.01

To supplement the survival analysis, we created a bar chart for the matched samples that shows the 12-month turnover rates by year of start date at DCP (see Figure 4.3). Naturally, all the start dates 2015 and earlier belong to the non-sim group, and all the start dates 2016 and

later belong to the sim group. A recent spike in 12-month turnover among new investigators is evident for those who started in 2019. Note that the 12 months following 2019 start dates coincide with the onset of the COVID-19 virus. The effects of the virus are one plausible explanation for the recent spike in turnover.

Figure 4.3

Turnover within 12 Months by Cohort (N=808)
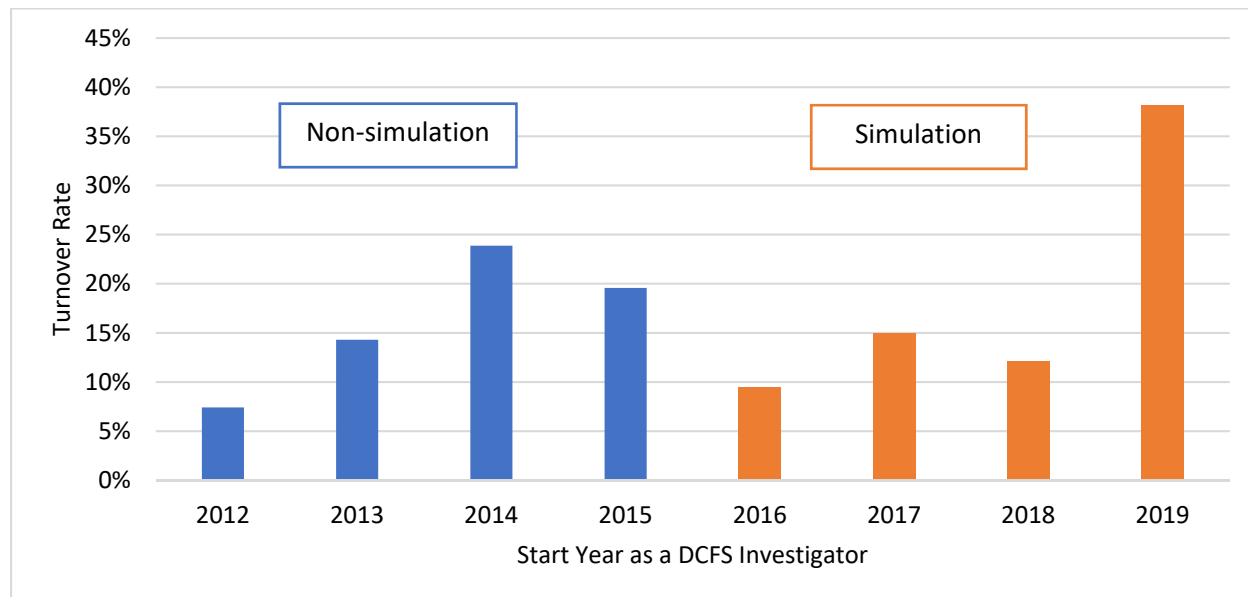
Discussion of the Survival Analysis

The results from the current survival analysis differ substantially from the survival analysis we conducted in FY2019. While in FY2019 report we found that sim-trained investigators were *less* likely to leave their jobs within the observation period, in the current analysis we found that sim-trained investigators were *more* likely to leave their jobs within the observation period. While several differences between the two analyses might contribute to the difference in findings, the most likely explanation is that the current analysis includes investigators who started at DCP in calendar year 2019, many within 12 months of the onset of the COVID-19 pandemic. We think the difference in the historical eras studied in the two analyses is likely to be the biggest explanation for the difference in results.

Discussion

We end our analysis with considerable ambiguity. Despite using the most sophisticated statistical methods we could use, we were unable to estimate clearly the effect of simulation training on turnover. There is no way, statistically or otherwise, to disentangle the effect of introducing simulation training from the effects of other historical change that coincided with simulation training. Note that the difference in historical era between sim and non-sim investigators encompasses a number of years. Though sim training occurs only in their first

year, comparing their employment over time meant that their first, second, third year and so forth occurred in very different years for the sim and non-sim groups.

It is possible that more data over time would help resolve this ambiguity. Over time, recent trends in employment such as the recent spike in turnover may become mere "blips" in longer term trends. If the trend in turnover over the course of a decade or more following 2016 looks very different than the trend prior to 2016, perhaps we can infer an effect of simulation training. However, DCFS is making efforts all the time to reduce turnover, so we should prepare to tolerate not knowing for certain what, if any, effect simulation training has had on turnover. In the future, additional states may implement simulation training of investigators, and perhaps we can learn more from comparing turnover between multiple states that have implemented simulation training and multiple states that have not.

Chapter 5: Analysis of the Relationships between Simulation Training and Re-Reports

One of the ultimate goals of the CPTA's simulation training program is to increase children's safety. This goal is included in CPTA's simulation training logic model.[26] Simulation training is thought to increase children's safety by helping investigators identify safety threats. If safety threats are identified, they can be addressed and children will be safer. If investigators are capable of identifying safety threats effectively, children judged to be safe would truly be safe. Therefore, it is relevant to compare sim-trained and non-sim-trained investigators on child safety.

Maltreatment re-report is a commonly used indicator to assess performance of a child welfare programs. For example, one of the Child and Family Service Review (CFSR) indicators is recurrence of maltreatment.[27] Several statewide child welfare program evaluations have included maltreatment re-report as an outcome, such as evaluations of the Illinois and Oregon differential response initiatives.[28]

However, a number of difficulties complicate the use of re-report as an outcome measure.[29] Many children who have been maltreated are not reported to child protective services, and many children who are reported have not been maltreated. Thus re-report is an imperfect measure of child maltreatment, and restricting re-report to *substantiated* or *indicated* does not solve the problem, because children who are truly at risk may have their investigations unsubstantiated because of lack of evidence.[30] There are other measurement issues with re-report as well;[31] for example, neglect is more likely to be re-reported than abuse because of the latter is event-based while the former represent a chronic insufficiency in care not typically represented in specific events.

A number of studies have found that families receiving ongoing services from child protective services agencies are actually *more* likely to experience a re-report than families who do not receive ongoing services.[32] One reason might be that families who receive ongoing services may be at greater risk than those who do not. However, Fuller and Nieto (2013) used propensity score matching to equate Illinois families who did and did not receive ongoing services on risk, and still found that families receiving child protective services were at greater risk than those

[26] Cross, T.P., Tittle, G. & Chiu, Y. (2017). *Program evaluation of simulation training for new DCFS investigators: Initial report.* Children and Family Research Center, School of Social Work, University of Illinois at Urbana-Champaign.
https://cfrc.illinois.edu/pubs/rp_20180131_ProgramEvaluationofChildProtectionTrainingAcademyforNewDCFSInvestigators:InitialReport.pdf

[27] https://capacity.childwelfare.gov/states/focus-areas/cqi/cfsr-data-syntax-toolkit/

[28] Fuller, T., Nieto, M., & Zhang, S. (2013). Differential response in Illinois: Final evaluation report. Urbana, IL: Children and Family Research Center, University of Illinois at Urbana-Champaign. Fuller, T., Braun, M., Chiu, Y.C., Cross, T., Nieto, M., Tittle, G., & Wakita, S. (2017). *Oregon differential response: Final evaluation report.* Urbana, IL: Children and Family Research Center, University of Illinois at Urbana-Champaign.

[29] Jenkins, B. Q., Tilbury, C., Mazerolle, P., & Hayes, H. (2017). The complexity of child protection recurrence: The case for a systems approach. *Child Abuse & Neglect*, *63*, 162-171.

[30] See Cross, T.P & Casanueva, C. (2009). Caseworker judgments and substantiation. *Child Maltreatment*, *14*, 38-52. Drake, B. (1996). Unraveling "unsubstantiated." *Child Maltreatment*, *1*, 261–271.

[31] Jenkins et al. (2017), et al.

[32] Jenkins et al. (2017), et al.

who did not.[33] Possibly Fuller and Nieto were not able to equate families who receive and do not receive child welfare services; Jenkins et al. (2017)[34] suggested that data taken from case records, which Fuller and Nieto used, may not adequately measure risk. Another possible explanation for families who receive ongoing child welfare services having higher risk is the possibility of a so-called surveillance effect—families who receive services may be observed more often and more closely than other families, increasing the likelihood of a maltreatment re-report. Chaffin and Bard's 2006 study found that surveillance effects were small, however.[35]

Despite the many studies that did not find positive effects of interventions on reducing re-reports, Barto and colleagues in 2018 found an ameliorative effect of the Massachusetts Child Trauma Project (MCTP) on re-reports. MCTP trained 71% of that state's child welfare workforce in trauma-informed case work practice and also increased access to evidence-based mental health services for trauma.[36] Children served by trainees of MCTP were significantly less likely to experience a re-report than children served by a comparison group of workers.

There is another problem with using re-report as an outcome measure in comparing cases at different historical periods. (here, the relevant historical periods are 2016 to the present, the era in which all new DCFS investigators have received simulation training; and prior to 2016, when no DCFS investigators received simulation training). There are historical trends in both reporting of maltreatment and re-reporting of child maltreatment. Illinois data show increases in recent years in both reports[37] and re-reports.[38] National data also show recent increases in report and re-reports.[39] There is considerable uncertainty about the causes of historical trends in child maltreatment reporting, but the causes are likely to relate to societal trends related to those factors that place children at risk or buffer them from risk.[40] We know of no research that specifically examines reasons for historical trends in re-reporting, but whatever factors drive reporting are likely to drive re-reporting as well. Using data from the Children and Family Research Center's Data Center,[41] we found a correlation of .93 between the number of

[33] Fuller, T. L., & Nieto, M. (2009). Substantiation and maltreatment rereporting: A propensity score analysis. *Child Maltreatment*, *14*(1), 27–37.

[34] Jenkins et al., ibid.

[35] Chaffin, M., & Bard, D. (2006). Impact of intervention surveillance bias on analyses of child welfare report outcomes. *Child maltreatment*, *11*(4), 301-312.

[36] Barto, B., Bartlett, J. D., Von Ende, A., Bodian, R., Norona, C. R., Griffin, J., Fraser, J. G., Kinniburgh, K., Spinazzola, J., Montagna, C., & Todd, M. (2018). The impact of a statewide trauma-informed child welfare initiative on children's permanency and maltreatment outcomes. *Child Abuse & Neglect, 81*, 149–160. https://doi.org/10.1016/j.chiabu.2018.04.023

[37] Illinois Department of Children and Family Services. (2021) *Six - Year Statistics on Child Protective Services. Data as of March 31, 2021*. Springfield, IL: Authors. https://www2.illinois.gov/dcfs/aboutus/newsandreports/Documents/ESS_Protective_Services.pdf

[38] Children and Family Research Center (2021). *Conditions of children in or at risk of foster care in Illinois: FY2021 Monitoring Report of the B.H. Consent Decree*. Urbana, IL: Children and Family Research Center, University of Illinois at Urbana-Champaign.

[39] Kim, H., & Maguire-Jack, K. (2021). Longitudinal changes in child maltreatment Reports. *American Journal of Orthopsychiatry*. Advance online publication. http://dx.doi.org/10.1037/ort0000565

[40] Finkelhor, D., & Jones, L. (2006). Why have child maltreatment and child victimization declined? *Journal of Social Issues*, *62*(4), 685-716.

[41] Children and Family Research Center (2021). *Child and Family Services Review (CFSR) Indicator Tables.* https://www.cfrc.illinois.edu/cfsr-tables.php?ind=pct_repeat_subreps

substantiated reports to DCFS from 2013 to 2019, and the percentage of those cases that had another substantiated re-report within 12 months. This suggest that the rate of re-reporting increases as the rate of reporting increases. Historical trends in re-reporting complicate the comparison of sim-trained and non-sim-trained investigators on re-reports on these cases, because sim-trained and non-sim-trained investigators began their work at DCFS in different historical periods.

To evaluate the relationship of simulation training to child safety, we compared sim-trained and non-sim-trained investigators on the likelihood that children in their investigations were involved in re-reports to DCFS. We took steps to try to circumvent the problem of historical trends confounding the comparison of sim-trained and non-sim-trained investigators.

Methods

Case Sample

Children and families receiving intact family services or involved in substitute care were not included, because the services they receive might affect the likelihood of a re-report. The case sample for this analysis thus included children or families **who did not receive any post-investigation services**.

Investigator Sample

Figure 5.1 below illustrates the two groups of investigators that were compared. The sim group included investigators who started to work at the Division of Child Protection (DCP) at DCFS between March 2016 and February 2018. The comparison group included investigators who started their work at DCP between January 2014 and December 2015.

Figure 5.1

Investigator Sample

Sim-Trained Group		Non-Sim Comparison Group
Investigators started at DCP between **March 2016 and February 2018**	Versus	Investigators started at DCP between **January 2014 and December 2015**
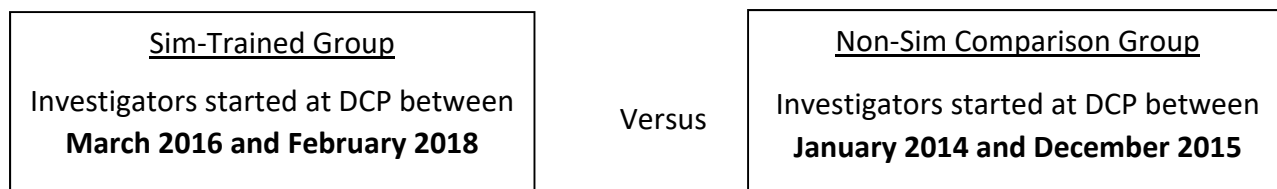
Data Collection

The proposed research had two data sources: DCFS employment data and data from DCFS' Statewide Automated Child Welfare Information System (SACWIS). CFRC already has access to the SACWIS data through its existing data sharing agreement with DCFS. The employment data file includes the following data fields: investigator ID (which can link with the investigation report information on SACWIS), DCFS hired date, DCP start date, DCP departure date, date of birth, gender, race, ethnicity, position when started at DCP, and education background.

Using DCFS SACWIS data, we tracked records of children or families seen by investigators to identify re-reports. We examined whether cases had a re-report within 6 months following the investigation.

Comparisons Between Sim and Non-Sim Investigators

We made two comparisons to examine the effect of simulation training on maltreatment re-reports (see Table 5.1). The first comparison contrasted sim-trained and non-sim-trained on investigations they worked on during the first two years in their career as an investigator.[42] Thus, in this comparison, the investigators had similar levels of experience. We call this the **Equivalent Experience Comparison**. This comparison had the disadvantage that sim and non-sim investigations being compared came from different historical periods (because all of the investigators hired after February 2016 had sim training, while none of the investigators hired before that had it). As we discuss above, there is historical change in re-reporting that may be correlated with historical change in reporting, and may be a function of change in societal factors underlying reports of abuse and neglect.

The second comparison contrasted sim-trained and non-sim-trained investigators on investigations during the same historical time period—we call this the **Equivalent Historical Period Comparison**. The second comparison had the disadvantage of comparing sim- and non-sim investigators with different levels of experience. Because sim-trained investigators were hired later, historically, than non-sim-trained investigators, they have less experience than non-sim-trained investigators in any given historical period. In the Equivalent Historical Period Comparison, sim-trained investigators had 0 to 2 years of experience in DCP while non-sim-trained investigators had 3 to 4 years of experience.

Table 5.1

Comparisons Between Sim-Trained and Non-Sim-Trained Investigators

Comparison	Investigator Group	DCP Start Date	Period in Investigators' Career	Date Range of Investigations Across Investigators
Equivalent Experience	Non-Sim	January 2014 to December 2015	Year One and Two	January 2014 to November 2017
	Sim	March 2016 to February 2018	Year One and Two	March 2016 to January 2020
Equivalent Historical Period	Non-Sim	January 2014 to December 2015	Year Three and Four	January 2016 to November 2019
	Sim	March 2016 to February 2018	Year One and Two	March 2016 to January 2020

---

[42] The investigations that were conducted on reports of child maltreatment in child-caring facilities were excluded because the investigation procedures for facilities is different from those developed for family settings (Illinois Department of Children and Family Services. (October, 2015). *Procedures 300.20 Reporting and Documenting Child Abuse or Neglect to the Department*. Springfield, IL: Author). Investigations that were conducted by supervisors were also excluded.

Variables

In this section, we discuss what variables were used in the analysis, including the dependent variable, the independent variable, and control variables. We were interested in including in the analysis a wide variety of variables that might influence the probability of a re-report, for the following reasons: a) a number of these variables might differ between sim and non-sim investigators, and confound the sim vs. non-sim comparisons, b) accounting for other variables that explain re-reporting increases the statistical power of the comparison between sim and non-sim investigators, by reducing the error term in tests of statistical significance, and c) accounting for other factors augments our understanding of re-reporting in general, helping us place the comparison of sim and non-sim on re-reporting in context. We benefited from a very large sample size, which enabled us to include a large number of variables in the analysis and still produce stable statistical models. More information on these variables is provided in Appendix B.

Maltreatment re-report. The dependent variable was defined as a substantiated maltreatment report within 6 months of investigation finding date.

Simulation training. This was a dichotomous variable indicating whether an investigator received or did not receive simulation training. As mentioned previously, the sim group included all DCFS investigators who started to work at DCP between March 2016 and February 2020. The non-sim group included DCFS investigators who started to work at DCP between January 2012 and December 2015. We omitted cases for investigators hired between January 2016 and February 2016, because we could not determine with precision whether or not they received simulation training.

Investigator control variables. We anticipated that different investigator characteristics might be related to investigator decision-making and thereby affect the probability of a re-report. We were also aware that sim and non-sim investigators differ on several characteristics, making it necessary to control for investigator characteristics in order to compare sim and non-sim investigators. We therefore included the following investigator characteristics in the analysis: investigator age at the time of hiring, investigator gender, investigator race/ethnicity, investigator education, investigator position at the time when they started at DCP (as an intern or child protection specialist), experience in DCFS prior to being hired as an investigator, and caseload at the time of investigation

Investigation characteristics. Characteristics of the original investigation might also influence the likelihood of a re-report. We included the following investigation-related variables in the analysis: Number of alleged victims, age of the youngest alleged victim, finding of the investigation, administrative region of the investigation, safety-decision of the last Child Endangerment Risk Assessment Protocol (CERAP)[43] conducted during investigation, number of

---

[43] The CERAP is a safety assessment protocol used in child protection investigations and child welfare services in Illinois. This "life-of-the case" protocol is designed to provide workers with a mechanism for quickly assessing the potential for moderate to severe harm to a child in the immediate or near future and for taking quick action to protect children. (References: Children and Family Research Center (2019). *Illinois Child Endangerment Risk Assessment Protocol: FY2019 Annual Evaluation.* Urbana, IL: Children and Family Research Center, University of

CERAP assessments during the investigation, number of domestic violence factors in the family, and number of substance abuse factors of adults in the family

Caregiver characteristics. Because caregiver characteristics may also affect the likelihood of a re-report, we included caregiver variables in the analysis. For each family, only one caregiver was selected. Since there might be more than one caregiver in the family, the caregiver was selected based on the following order of priority: mother, father, other relative non foster, relative foster parent, nonrelative foster parent, group home or residential, child daycare provider, unmarried partner of parent, other professionals. The following caregiver variables were used: age, gender, race, whether the caregiver was a parent, number of prior substantiated CPS reports, allegation type in prior substantiated CPS report(s) (physical abuse, sexual abuse, substance exposure, emotional abuse, lack of supervision, environmental neglect, substantial risk of harm, other neglect).

Analytic Approach

Cox's proportional hazard regression is a type of event history analysis that examines the occurrence of a particular event (in this case, a substantiated maltreatment report) over time. This method was used to assess the relationship of investigator simulation training to a substantiated maltreatment re-report in the investigator's cases. For each investigation, the observation period started on the date at which there was a finding in the investigation (the finding date) and continued until either a substantiated maltreatment report occurred or the 6-month follow-up period ended. Because the unit of analysis was the investigation and each investigator could conduct multiple investigations, it was necessary to adjust the analyses to account for the fact that observations were not independent of one another. Therefore, the Cox's regression procedure was run with the COVSANDWICH variance adjustment, which uses the robust test described by Lin and Wei[44] to evaluate the effects of the predictor variables in the model. To avoid losing cases because of missing data, "missingness" of information on several variables was treated as a variable. This enabled us to include more cases in the analysis while controlling for the potentially biasing effect of missing data. All analyses in this chapter were completed using SAS v9.4.

Results

Equivalent Experience Comparison: Investigations within 24 months of hiring date

For the Equivalent Experience Comparison, there were 270 investigators in the simulation group and 134 investigators in the non-simulation group. The simulation group conducted 49,451 investigations while the non-sim group conducted 23,637 investigations.

Table 5.2 shows the results of the survival analysis for the Equivalent Experience Comparison, including the statistical significance and the hazard ratio (HR) for each variable. Statistical

---

Illinois at Urbana-Champaign. Illinois Department of Children and Family Services. (May, 2013). *Procedures 300 Appendix G –Child Endangerment Risk Assessment Protocol.* Springfield, IL: Author).

[44] Lin, D., & Wei, L.J. (1989). The robust inference for the Cox proportional hazard model. *Journal of the American Statistical Association, 84*, 1074-1078.

significance has serious limitations as a measure of importance,[45] especially when sample sizes are very large, as in the current analysis. When sample sizes are very large, even trivial statistical effects can be statistically significant.

We focus instead on the hazard ratio (HR), which is a measure of effect size—an assessment of the relationship between a variable and re-report. In this analysis, the HR is the ratio of the probability of an event (in this case, the probability of a substantiated maltreatment report) in the one group (let us call it Group A) divided by the probability of the same event in the other group (Group B). If the probability of the event is similar in the two groups, the HR will be close to 1. An HR greater than 1 indicates that the event is *more likely* to occur in Group A than in the Group B, while an HR less than 1 indicates that the event is *less likely* to occur in Group A than the Group B. For example, an HR of 2 indicates that the event is twice as likely (or 200% as likely) to occur in Group A than in the comparison group, while an HR of .5 indicates that the event is half as likely to occur Group A as in Group B. For continuous variables (such as the number of alleged victims), the HR indicates the change in risk associated with a one unit increase in the predictor variable. Azuero (2016) provided guidance about interpreting the size of the HR.[46] For a comparison of two groups, he suggests that a hazards ratio of 1.3 represents a small effect, 1.9 a medium effect, and 2.8 a large effect. Writing as a medical researcher, Azuero suggests an HR of 1.57 as a lower limit of an effect that is likely to be clinically relevant. The hazard ratios at these levels are somewhat less when the predictor is a continuous variable: 1.14 for a small effect, 1.47 for a medium effect, and 1.9 for a large effect, and 1.3 for an effect likely to be clinically relevant.[47]

Because the sample size of our study was very large, a number of variables had a statistically significant relationship to re-report, but a smaller set of variables had effects that met Azuero's threshold for clinical relevance. When the original investigation led to a substantiated finding, a substantiated re-report was 1.62 times more likely to occur within 6 months than when the original investigation was unfounded. Substantiated re-reports were 2.26 times more likely when the caregiver was a parent compared to cases in which a non-parent was the caregiver.

Several variables stemming from DCFS' assessment of abuse and neglect in the original investigation predicted re-reports. Re-reports were 1.84 more likely when there was substance exposure, 2.19 more likely when there was lack of supervision, 2.11 more likely when there was environmental neglect, and 1.55 times more likely when there was other neglect. All of these hazard ratios were comparisons between a case having this alleged type of maltreatment vs. not having it. We did not include the number of types of alleged maltreatment as a variable because it was correlated with each specific type, but the significant relationships of all these types with re-report suggests that the number of types of alleged maltreatment is also related to the risk of re-report. Re-reporting was also 1.97 times more likely when the investigator in the original investigation judged that there was a substantial risk of harm. A number of

---

[45] Cohen, J. (1994). The earth is round (p<.05). *American psychologist, 49,(*12), 997-1003.

[46] Azuero, A. (2016). A note on the magnitude of hazard ratios. *Cancer*. 1298-1299.

[47] For two-group comparison hazard ratios less than 1.0, a small effect = .77, a medium effect = .53, a large effect = .36, and a clinically relevant effect = .64. For continuous variable hazard ratios less than 1.0, a small effect = .88, a medium effect = .68, a large effect = .53, and a clinically relevant effect = .77.

different missing data variables were related to re-report as well—cases with a re-report have more valid information and so were less likely to have missing data on our variables (see Table 5.2).

The HR for simulation training was 1.32, a statistically significant but small effect that was below Azuero's threshold for clinical relevance (Table 5.2). A re-report was more likely when a sim investigator investigated the original allegation, though the difference was small. Note that there was also a small but statistically significant relationship of year of original investigation to re-report. Re-reports were more likely in the later years represented in the sample than in the earlier year. Because all the original investigations in the Equivalent Experience Comparison sample in 2014 and 2015 were investigated by non-sim investigators, and all the cases in 2018 thru 2020 were investigated by sim-trained investigators, year is confounded with the sim vs. non-sim comparison, and it is impossible to disentangle the two variables fully. It is impossible to rule out year of investigation as an explanation for the difference between the sim and non-sim groups on re-report.

Table 5.2

Equivalent Experience Comparison: Survival analysis predicting maltreatment re-report for investigations within 2 years of hiring date (n = 73,088)

Variable	Parameter estimate	SE	Chi-Square	p	Hazard Ratio
Year of investigation occurred	0.07	0.02	9.64	0.0019	1.08
Investigation finding					
Compare substantiated with unfounded findings	0.48	0.04	118.14	<.0001	1.62
Compare missing with non-missing findings	-0.05	0.22	0.05	0.8254	0.95
Administrative region (reference group=Cook)					
Northern	-0.06	0.07	0.72	0.3977	0.94
Central	0.16	0.07	4.40	0.0359	1.17
Southern	0.06	0.08	0.53	0.4656	1.06
Whether parent was the caregiver	0.81	0.20	17.35	<.0001	2.26
Whether the age of caregiver was missing	-0.87	0.26	11.49	0.0007	0.42
Caregiver age	-0.01	0.00	5.68	0.0172	0.99
Caregiver gender					
Valid vs. missing data	0.31	0.09	12.76	0.0004	1.36
Compare female with male	0.08	0.03	5.43	0.0198	1.08
Caregiver race/ethnicity (reference group=all others (non-missing))					
non-Hispanic Black	0.09	0.08	1.32	0.25	1.10
non-Hispanic White	0.12	0.07	2.62	0.1058	1.13
Hispanic	-0.01	0.09	0.01	0.905	0.99
Other	-0.18	0.20	0.77	0.3788	0.84
Total number of alleged victims in the household	0.11	0.02	32.06	<.0001	1.12

Number of CERAPs during investigation (reference group=1)					
CERAP count=2	-0.04	0.04	0.87	0.3512	0.96
CERAP count=3	0.07	0.08	0.70	0.4029	1.07
Safety decisions of last CERAP					
Compare missing with non-missing	-0.21	0.18	1.38	0.2394	0.81
Compare safe with unsafe	0.03	0.06	0.25	0.6183	1.03
Total number of domestic violence factors in the Family	0.01	0.00	3.91	0.0481	1.01
Whether domestic violence screening forms were missing	-0.09	0.23	0.16	0.6848	0.91
Total number of substance abuse factors in the Family	0.06	0.01	90.81	<.0001	1.06
Whether substance abuse screening forms were missing	0.16	0.22	0.52	0.4705	1.17
Age of the youngest alleged victim	-0.05	0.00	105.53	<.0001	0.95
Simulation training	0.27	0.07	14.34	0.0002	1.32
Investigator age at Hire	0.00	0.00	0.03	0.8644	1.00
Investigator gender=Female	-0.01	0.06	0.05	0.8262	0.99
Race/Ethnicity of Investigator (reference group=non-Hispanic White)					
Non-Hispanic Black	-0.07	0.05	1.47	0.2255	0.94
Hispanics	-0.12	0.08	2.05	0.152	0.89
Other	-0.09	0.16	0.30	0.5821	0.91
Prior experience in DCFS before DCP initial start date (in month)	0.00	0.00	0.95	0.3285	1.00
Whether the investigators had a Social Work degree	0.02	0.05	0.15	0.7017	1.02
Investigator education (reference group=Bachelor's degree)					
Higher school or associate degree	-0.17	0.16	1.11	0.2925	0.84
Master's degree or higher	-0.07	0.07	0.97	0.3236	0.93
Investigator position at DCP initial start date: Children and family service intern	-0.05	0.06	0.77	0.3797	0.95
Investigator caseload during the investigation	0.00	0.00	1.16	0.2817	1.00
Total number of prior substantiated CPS reports of the family	0.10	0.01	64.36	<.0001	1.11
Prior substantiated allegation type					
Sexual abuse	-0.11	0.15	0.51	0.475	0.90
Physical abuse	0.21	0.09	5.27	0.0217	1.24
Substance exposure	0.61	0.13	23.07	<.0001	1.84
Emotional abuse	0.48	0.36	1.78	0.1822	1.62
Lack of supervision	0.79	0.06	168.29	<.0001	2.19
Environmental neglect	0.74	0.09	63.50	<.0001	2.11
Other neglect	0.44	0.11	16.20	<.0001	1.55

Substantial risk of harm	0.68	0.05	186.58	<.0001	1.97

Equivalent Historical Period Comparison: Investigations that happened during the same time period

For the Equivalent Historical Period Comparison, there were 270 investigators in the simulation group and 88 investigators in the non-simulation group. The sim group conducted 49,451 investigations while non-sim group conducted 15,994 investigations. Results from survival analysis are presented in Table 5.3.

In most ways, the results for the Equivalent Historical Period Comparison paralleled those of the Equivalent Experience Comparison. When the original investigation led to a substantiated finding, a substantiated re-report was 1.35 times more likely to occur within 6 months than when the original investigation was unfounded. Substantiated re-reports were 2.97 times more likely when the caregiver was a parent compared to cases in which a non-parent was the caregiver. Each of the following types of alleged maltreatment was moderately to substantially related to re-reporting substance exposure (1.63), emotional abuse (2.32), lack of supervision (1.97), environmental neglect (1.77) and other neglect (1.50) (Table 5.3). The same was true of substantial risk of harm (1.94). The results for the missing data variables were similar to the other comparison (Table 5.2).

However, in the Equivalent Historical Period Comparison, sim training was not significantly related to the likelihood of re-report. In other words, when we compared the sim group and non-sim group during the same years in history (even though their experience differed), the sim and non-sim cases were about equally likely to have a re-report.

Table 5.3

Equivalent Historical Period Comparison: Survival analysis predicting maltreatment re-report for comparing investigations that were undertaken during the same time period (n = 65,445)

Variable	Parameter estimate	SE	Chi-Square	p	Hazard Ratio
Year of investigation occurred	0.08	0.02	9.75	0.0018	1.08
Investigation finding					
Compare substantiated with unfounded findings	0.30	0.05	42.07	<.0001	1.35
Compare missing with non-missing findings	-0.51	0.28	3.36	0.0667	0.60
Administrative region (reference group=Cook)					
Northern	-0.03	0.07	0.23	0.6298	0.97
Central	0.20	0.08	7.09	0.0078	1.22
Southern	0.17	0.08	4.75	0.0293	1.19
Whether parent was the caregiver	1.09	0.21	27.11	<.0001	2.97
Whether the age of caregiver was missing	-0.71	0.24	8.92	0.0028	0.49
Caregiver age	-0.01	0.00	6.31	0.012	0.99

Caregiver gender					
Compare missing with non-missing	0.32	0.08	15.10	0.0001	1.38
Compare female with male	0.07	0.03	4.85	0.0277	1.08
Caregiver race/ethnicity					
non-Hispanic Black	0.13	0.08	2.39	0.1224	1.14
non-Hispanic White	0.15	0.08	3.78	0.0518	1.16
Hispanic	0.02	0.09	0.06	0.8104	1.02
Other	-0.15	0.21	0.51	0.4756	0.86
Total number of alleged victims in the household	0.11	0.02	31.12	<.0001	1.11
Number of CERAPs during investigation (reference group=1)					
CERAP count=2	-0.06	0.04	1.67	0.1964	0.95
CERAP count=3	0.04	0.08	0.19	0.6627	1.04
Safety decisions of last CERAP					
Compare missing with non-missing	-0.17	0.16	1.21	0.2713	0.84
Compare safe with unsafe	-0.01	0.07	0.03	0.8575	0.99
Total number of domestic violence factors in the Family	0.01	0.00	7.15	0.0075	1.01
Whether domestic violence screening forms were missing	-0.07	0.27	0.07	0.7885	0.93
Total number of SA factors in the Family	0.06	0.01	103.25	<.0001	1.06
Whether SA screening forms were missing	0.11	0.26	0.20	0.6551	1.12
Age of the youngest alleged victim	-0.05	0.00	89.00	<.0001	0.96
Simulation training	-0.01	0.06	0.02	0.8904	0.99
Investigator age at Hire	0.00	0.00	0.02	0.8897	1.00
Investigator gender=Female	-0.03	0.06	0.28	0.5984	0.97
Race/Ethnicity of Investigator (reference group=non-Hispanic White)					
Non-Hispanic Black	-0.07	0.06	1.58	0.2081	0.93
Hispanics	-0.16	0.08	3.37	0.0664	0.86
Other	0.07	0.10	0.46	0.4957	1.07
Prior experience in DCFS before DCP initial start date (in month)	0.00	0.00	0.01	0.9423	1.00
Whether investigators have a Social Work degree	0.01	0.05	0.05	0.8186	1.01
Investigator education (reference group=Bachelor's degree)					
Master's degree or higher	-0.07	0.08	0.76	0.3829	0.93
Higher school or associate degree	-0.11	0.16	0.48	0.4902	0.90
Investigator position at DCP initial start date: Children and family service intern	-0.04	0.07	0.33	0.5652	0.96
Investigator caseload during the investigation	0.00	0.00	2.55	0.11	1.01
Total number of prior indicated CPS reports of the family	0.11	0.01	65.38	<.0001	1.11

Prior indicated allegation type					
Sexual abuse	-0.22	0.16	1.83	0.1756	0.80
Physical abuse	0.19	0.09	4.01	0.0452	1.21
Substance exposure	0.49	0.13	14.56	0.0001	1.63
Emotional abuse	0.84	0.32	7.04	0.008	2.32
Lack of supervision	0.68	0.06	118.82	<.0001	1.97
Environment neglect	0.57	0.10	33.94	<.0001	1.77
Other neglect	0.40	0.11	14.29	0.0002	1.50
Substantial risk of harm	0.66	0.05	182.70	<.0001	1.94

Discussion

The results of the analyses in these chapter do not provide evidence for an effect of simulation training on re-reports. In the Equivalent Experience Comparison, investigation by sim investigators were slightly more likely to result in a re-report, but the difference was small and cannot be clearly separated from historical change in re-report rates. There was no difference in the Equivalent Historical Period Comparison, even though the sim group in that comparison was less experienced than the non-sim group. Thus we do not have evidence to support the supposition in the CPTA simulation training logic model that simulation training reduces re-reporting.

These results need to be put in context. Research has shown that those who receive ongoing child welfare services are *more* likely to be re-reported, so standard interventions may not reduce risk of re-report. Investigation is a brief intervention, and no matter how well conducted, it may not be powerful enough to have an impact on re-reports. Barto and colleagues' research cited above suggests that a substantial trauma-informed intervention may reduce the risk of re-reporting, but this is a fairly substantial intervention specifically targeted to factors that place families at risk for re-reporting.

Perhaps re-reporting should be removed as an outcome from the CPTA simulation training logic model. Another option is to consider this outcome aspirational and consider what impact improved investigation may have in a society in which families can transition from an effective and compassionate investigation to receiving evidence-based interventions closely targeted to those factors that are placing families at risk for re-reports.

Chapter 6: Conclusion

FY2021 required endurance and ingenuity to sustain what was good. The COVID-19 pandemic that erupted in the Spring of 2019 persisted throughout the entire fiscal year and restrictions due to the pandemic are still in place now that the fiscal year is ending. The Child Protection Training Academy had to stop in-person training, and, after a gap of several months, moved entirely to virtual training in August 2020. Most of the data analyzed in this report come from the virtual trainings. This provided a challenging but perhaps useful test of the adaptability of simulation training of child protection workers, providing lessons about options for delivering the training in the future. An overarching lesson that should not be lost as we explore the findings of the data is that simulation was successfully adapted and investigators for DCFS received this experience despite the many barriers.

Assessing Simulation Training Delivered Virtually

The program evaluation provides data supporting the value of simulation training even when delivered virtually. Trainees' reports from the Daily Experience of Simulation Training (DEST) measure showed increases in confidence in child protection skills during the virtual simulation training weeks, with confidence levels and upward trends similar to those found in the past with in-person training. On the DEST, a majority of trainees rated the feedback as very helpful and the vast majority as at least helpful, and on average they rated the debriefing as very effective. On the post-training survey, a majority of respondents agreed or strongly agreed with the items indicating satisfaction with simulation training. Apart from ratings from two trainings in which several trainees had a challenging experience (see Section 6.2 below), these ratings were comparable to those of previous years. In open-ended comments, a number of people expressed appreciation for simulation training and reported meaningful learning from it. The CPTA achieved these outcomes despite the disruption caused by the COVID-19 pandemic and the need to adapt virtual methods quickly to deliver experiential learning, which is built around in-person scenarios.

Yet some results suggest that trainees did not respond as positively to the virtual simulation training as previous trainees responded to in-person simulation training. In some ways this is expected given the stresses of the pandemic in general and on the training itself to quickly adapt and move forward. In open-ended comments, a number of respondents volunteered the opinion that there were limitations in the online training because of its virtual nature and felt that in-person training would have been better. It is also worth considering whether the occurrence of trainings in which several trainees had challenging experiences is related to any extra difficulties due to conducting the trainings virtually. As we discussed in Chapter 3, data from this fiscal year suggest that online training can be a helpful adjunct when in-person training is impossible or impractical, but it is probably not wise to consider it equivalent to in-person simulation training.

Trainees Who Found Simulation Training Challenging

Some respondents reported challenging experiences in simulation training. Most of these were concentrated in two different training cohorts. Each of these cohorts had several respondents who used the open-ended items in the post-training survey to describe their challenging

experience. In last year's final program evaluation report, we reported that a small percentage of simulation trainees who completed the post-training survey had a challenging experience, though the vast majority reported positive experiences. One question is why there was evidence of trainees with challenging experiences in the post-training survey data but little or no such evidence in the DEST data.

As we discussed in Chapter 3, simulation training is an emotionally powerful experience, and the fact that some people experience difficulties with it should be an expected outcome in both simulation and child welfare work in the field. Child protection investigators' own personal experience of trauma is likely to be a factor. Participants' perception provides an opportunity to consider adjustments in training that could decrease challenges for participants. The existing trauma-informed framework offers a model for continuous quality improvement in this specific area.

This year's findings make us consider the role group process can play in simulation training. Considerable evidence this year and previous years suggests that most training cohorts experience a supportive group process that promotes learning. This year's results suggest that the group process can be challenging on occasion and multiple trainees in a given group can have a challenging experience at the same time. There may be a greater risk of this when training is virtual. We suggest considering additional steps to monitor and support the group process. Classroom trainers could attend simulation training to provide support to the trainees, who they have worked with for four weeks (classroom trainers attended in-person simulation training in the past but did not attend the virtual trainings in FY2021). Items might be added to the DEST that are more attuned to trainees' emotional experience. Perhaps end-of-the-day group briefings could include more time for trainees to offer feedback on their experience. A structured peer review process might be added to trainings. Trainers and supervisors could build in a regular mid-training review process. CPTA has engaged in a continuous quality improvement process over more than five years and has had success in innovating to enhance the training experience. One simple step is to access the post-training survey data regularly throughout the fiscal year to promote a timely response to any trainings in which multiple trainees have challenging experiences.

More research is needed regarding trainees who have challenging simulation training experiences. Several questions deserve research. What methods are effective for identifying in real time trainees having difficult experiences? What interventions can improve trainees' experience if at first it is difficult for them? What interventions could support trainees post-training if they have a challenging simulation training experience? If struggling with simulation training can be an indicator that child protection work is a poor fit to a trainee's capabilities, we need to be able to identify when that happens and what can be done either to improve that fit or minimize the impact of turnover for DCFS.

Culture and Diversity

On the post-training survey, one trainee this year mentioned the need to focus more on trainers' cultural competence, and trainees last year expressed the wish for greater diversity among the scenarios and actors used. These comments suggest the value of examining more the cultural competence of the program and exploring ways of addressing diversity more fully

in the training. Some research suggests the existence of racial bias in child welfare decision-making suggesting the need to deal with this issue more in training.[48]

Evaluating Trainee Behavior

Training is designed to increase trainees' learning as a means to influence trainee behavior and thereby produce better results. Recognizing this, Kirkpatrick and Kirkpatrick's manual on evaluating training posits four levels of evaluation: Level 1—Reaction, Level 2—Learning, Level 3—Behavior, Level 4—Results.[49] Applying this to simulation training, Level 1 involves assessing trainees' reaction to the training, Level 2 involves assessing what trainees learn, Level 3 involves assessing how training relates to trainees' behavior on the job, and Level 4 involves assessing how simulation training benefits DCFS clients. Data on the upper levels increase the value of program evaluation of training but are more difficult to obtain.

The analysis of turnover in Chapter 4 evaluates simulation training in terms of an important trainee behavior: leaving their job as investigators. This is a critical variable, because high turnover leads to increased inexperience and increased cost for the agency as well as negative impact for children and families. But turnover is very difficult to study as an effect of simulation training, because there are so many factors in addition to training that could influence turnover. Unsurprisingly, we found no relationship between simulation training and turnover given the many different factors that could influence turnover. One recent factor that likely affected turnover was the COVID-19 pandemic. Moreover, it is difficult to estimate the effect of simulation training on turnover because of limitations in what we can use as a comparison group.

As desirable as it is to measure investigators' behavior as a way to evaluate simulation training, it is difficult. One method would be to gather data on investigators on the job. We know of no previous research that does this. We can however extrapolate from studies in related fields and imagine a study that might successfully gather data on investigators. One could imagine videotaping child protection investigations, and rating the quality of the interaction with families, as has been done in research on family therapy.[50] Another pathway could be having qualified observers accompany investigators to and rating their behavior during actual investigations . One could also imagine interviewing families about their experience with child protection investigators. Fuller and colleagues at the Children and Family Research Center interviewed caregivers about their satisfaction with a differential response (DR) intervention by DCFS.[51] Since families receiving DR have been reported because of suspicion of child maltreatment, it seems reasonable to extend Fuller and colleagues' methods to families

---

[48] See, e.g., Dettlaff, A. J., Rivaux, S. L., Baumann, D. J., Fluke, J. D., Rycraft, J. R., & James, J. (2011). Disentangling substantiation: The influence of race, income, and risk on the substantiation decision in child welfare. *Children and Youth Services Review*, *33*, 1630–1637.

[49] Kirkpatrick, D.L. & Kirkpatrick, J.D. (2006). *Evaluating Training Programs: The Four Levels*. Third Edition. San Francisco: Berrett-Koehler Publishers, Inc.

[50] Cervantes Camacho, V., Mancini, T., Zaccaria, C., & Fruggeri, L. (2020). Testing the use of the System for Observing Family Therapy Alliances (SOFTA) in audio-recorded therapeutic sessions. *Couple and Family Psychology: Research and Practice*, *9*(2), 90–99.

[51] Fuller, T., Nieto, M. & Zhang, S. (2013). *Differential Response in Illinois: Final Evaluation Report*. Children and Family Research Center, School of Social Work, University of Illinois at Urbana-Champaign.

involved in investigations. Of course, even if interviews revealed that families had positive appraisals of sim-trained investigators, without an adequate comparison group, we could not infer that simulation training caused this outcome. As we have seen, there are limitations in using non-sim trained DCFS investigators as a comparison group, and the number of non-sim trained DCFS investigators is diminishing. Nevertheless, this would be pioneering research that could expand our knowledge about simulation training while also increasing our basic knowledge about investigator behavior and families' experience during investigations.

A promising approach to assessing trainee behavior is to measure the competency of trainees who participate in simulation training. Havig and colleagues developed a reliable measure of rating simulation trainees on three skill dimensions: rapport-building, communication and information-gathering, and safety assessment and ending.[52] CPTA has recently embarked on a plan to adapt this method to assess the competency of its trainees. One question to address in developing this for CPTA is whether trainers would rate trainees themselves or qualified observers would rate trainees, either through live observation or by watching video recordings. Havig et al. had observers rate trainees from video recordings. Having trainers rate trainees may be easier to implement and may be less likely to raise trainees' anxiety, but it is difficult to eliminate the possibility of trainer bias, particularly since the trainers' have a stake in the ratings. On the other hand, having outside observers rate trainee behavior might increase trainees' anxiety more, which might run counter to CPTA's goal of creating an environment for trainees that is as safe as possible. One must also acknowledge a limitation of rating trainees' behavior in the simulations: one cannot necessarily extrapolate the results to trainees' behavior in the field.

Evaluating Client Results

We examined the relationship simulation training and re-reporting based on the logic model for simulation training. The re-report analysis in Chapter 5 did not find that cases of investigators who had simulation training were less likely to have re-reports than cases of investigators who did not have simulation training. Given previous research that shows a negative effect of providing child protective services on re-reporting, and historical trends toward increased in re-reporting in recent years, it was probably unrealistic to expect simulation training to show an effect on re-reporting. It would be valuable to show the relationship between simulation training and client results, because the only purpose ultimately of simulation training is to benefit children and families. However, it is difficult to show this relationship, given the limited degree to which we can measure client outcomes, the variety of factors that can influence client outcomes, and historical trends that could influence any client outcome. Perhaps client's experience of the investigation would be the most important outcome to measure in relation to simulation training, because the most important lesson of simulation training may be learning to engage clients and maintain the relationship throughout the process of investigation. The same type of interview we discussed in the last section could be used not only to gather the

---

[52] Havig, K., Pharris, A., McLeod, D.A, Natale, A.P., & Miller-Cribbs, J. (2020) Assessing new child welfare worker competency through social simulation with standardized clients: rubric development and pilot testing, *Journal of Public Child Welfare*, *14*(5), 531-552.

client's appraisal of the investigator but also to measure the client's feelings about the investigation experience.

Final Words

Simulation is a rapidly growing method for training child protective services workers, but it is still early in its development and much remains to be learned through research and program evaluation. Results from the program evaluation of CPTA over a five-year period indicates that simulation training is a powerful and promising method in child welfare that deserves further development and study. One challenge for doing this is the paucity of research both on simulation training in child welfare and on training in general in child welfare and child welfare worker competency and performance. The lack of research on child protection investigators is particularly challenging This program evaluation represents a critical step in the continuous improvement of the Illinois DCFS program as well as a contribution to the growing body of knowledge about the value of simulation training for child welfare professionals. Having no established research methods and little established knowledge to rely on provides the opportunity to be pioneers in our search for knowledge and provide leadership to the field

The findings of this year's program evaluation suggest that studies examining field outcomes such as turnover and re-report are not likely to be fruitful at this stage, while the immediate impact of simulation training needs to be understood better. Implementing a system of rating trainees on rubrics, as CPTA is planning to do, could generate substantial knowledge about how simulation training is having an effect on specific skills and what specific elements of training that could be improved. Combined with the development of a measure of problem-based learning (see Chapter 1), this will significantly upgrade data that we can use to examine the immediate impact of simulation training.

Appendix A: Content of Daily Experience of Simulation Training 3.0

□ Day 1 Morning (Baseline) □ Day 1 Afternoon □ Day 2 □ Day 3 □ Day 4 □ Day 5

- At which site are you taking the training? □ Chicago□ Springfield
  1. Before you came to simulation training:
  2. When did you complete the foundational **classroom** training? __Month;__Year (Dropdowns)
  3. How much time did you spend in on job training (OJT)? □ none □ less than 1 week □ 1-2 weeks □ 3-4 weeks □ 5-6 weeks □ 7-8 weeks □ more than 8 weeks.
  4. What tasks did you do during OJT? (check that all apply) □ read related documents, such as case files or procedures □ learned about SACWIS □ shadowed seasoned investigators □ worked on investigation reports □other, specify_____

- With (1) being lowest and (7) being highest, please check the appropriate number to indicate your level of confidence in the following skill areas TODAY.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Low			Moderate			High
Gather info from collateral contacts	○	○	○	○	○	○	○
Think critically on facts vs. hypotheses	○	○	○	○	○	○	○
Engage families	○	○	○	○	○	○	○
Assess safety	○	○	○	○	○	○	○
Integrate compassion and investigative skill	○	○	○	○	○	○	○
Address any concerns about family statements and behaviors	○	○	○	○	○	○	○
Identify family strengths	○	○	○	○	○	○	○
Explain need for safety plan and/or protective custody	○	○	○	○	○	○	○
Explain DCFS role and expectations for keeping children safe	○	○	○	○	○	○	○
Answer pointed questions from parents and caregivers	○	○	○	○	○	○	○
Address underlying conditions such as domestic violence, substance abuse,	○	○	○	○	○	○	○

mental health, developmental disabilities							
Testify in court	○	○	○	○	○	○	○
Work as a DCFS investigator	○	○	○	○	○	○	○

- With (1) being lowest and (7) being highest, please check the appropriate number to indicate **the debriefing TODAY.**

	(1) Extremely ineffective	(2)	(3)	(4) Somewhat effective	(5)	(6)	(7) Extremely effective
Debriefing identified the areas in which I need to grow.							
In-class group debriefing	○	○	○	○	○	○	○
Individual debriefing	○	○	○	○	○	○	○
Debriefing provoked in-depth discussion that led me to reflect on my skills.							
In-class group debriefing	○	○	○	○	○	○	○
Individual debriefing	○	○	○	○	○	○	○
Debriefing allowed me to connect with class materials and their practical application.							
In-class group debriefing	○	○	○	○	○	○	○
Individual debriefing	○	○	○	○	○	○	○

- Please answer the following questions regarding the feedback that you received in **today's training**:

	very unhelpful	unhelpful	helpful	very helpful	N/A
I found the simulation facilitator's feedback to be…	○	○	○	○	○
I found the feedback of **foundational trainer** who assisted with the simulation facilitator to be…	○	○	○	○	○
I found the feedback of "family members" in general	○	○	○	○	○
I found the feedback of "medical professionals" in	○	○	○	○	○

general

| I found the feedback of "courtroom professionals" in general | ○ | ○ | ○ | ○ | ○ |

Today's reflective log:

- What was the most helpful feedback that you learned from your individual debriefing? And why?

- What were the most meaningful concepts or skills you learned from your classroom/foundation training so far? / What were the most meaningful concepts or skills you learned today?

Appendix B : List of Variables in Maltreatment Re-report Study

Variable Name	Description	Value
Calendar Year	Year of investigation occurred	2014-2020
Finding of Investigation 1	Compare substantiated with unfounded findings	1=Substantiated; -1=Unfounded 0=Missing/Pending
Finding of Investigation 2	Compare missing with non-missing findings	2=Missing/Pending; -1=Substantiated or Unfounded
Region: Northern	Region of investigation assigned: Northern	0=No; 1=Yes Reference group=Cook
Region: Central	Region of investigation assigned: Central	0=No; 1=Yes Reference group=Cook
Region: Southern	Region of investigation assigned: Southern	0=No; 1=Yes Reference group=Cook
Role of caregiver: Parent(s)	Whether parent was the caregiver	1=parent; 0=Non-parent
Caregiver age	Age of caregiver	Continuous (missing age was replaced by the mean)
Caregiver age: missing	Whether the age of caregiver was missing	1=Missing; 0=Non-missing
Caregiver Gender 1	Compare missing with non-missing	2=missing; -1=female or male
Caregiver Gender 2	Compare female with male	-1=male; 1=female 0=missing
Caregiver Race: Black	Race/Ethnicity of caregiver: non-Hispanic Black	1=Black; 0=White/Hispanic/other -1=missing
Caregiver Race: White	Race/Ethnicity of caregiver: non-Hispanic White	1=White; 0=Black/Hispanic/other -1=missing
Caregiver Race: Hispanic	Race/Ethnicity of caregiver: Hispanic origin	1=Hispanic; 0=White/Black/other -1=missing
Caregiver Race: Other	Race/Ethnicity of caregiver: other	1=other; 0=White/Black/Hispanic -1=missing

Number of alleged victims	Total number of alleged victims in the household	1 to 4; 4 includes 4 and more children
CERAP count=2	Total number of CERAPs done during the investigation =2	1=2 CERAP assessments 0=1 CERP assessment
CERAP count=3	Total number of CERAPs done during the investigation =3	1=3 or more CERAP assessments 0=1 CERP assessment
Safety Decision of Last CERAP 1	Compare missing with non-missing	2=missing; -1=safe/unsafe
Safety Decision of Last CERAP 2	Compare safe with unsafe	-1=unsafe; 1=safe 0=missing
Number of Domestic Violence Factors	Total number of domestic violence factors in the Family	Continuous; If it is missing, replace with the mean score
Missing Domestic Violence Assessment	Whether domestic violence screening forms were missing	0=No; 1=Yes
Number of Substance Abuse Factors	Total number of domestic violence factors in the Family	Continuous; If it is missing, replace with the mean score
Missing Substance Abuse Assessment	Whether domestic violence screening forms were missing	0=No; 1=Yes
Age of the youngest alleged victim		continuous
Simulation training	Whether investigators received a simulation training	0=Non-sim; 1=Sim
Investigator Age at Hire	Investigator's age at the Division of Child Protection (DCP) initial start date	continuous
Investigator Gender		0=Male; 1=Female
Investigator Race: Black	Race/Ethnicity of Investigators: Non-Hispanic Black	0=No; 1=Yes Reference group=White
Investigator Race: Hispanics	Race/Ethnicity of Investigators: Hispanics	0=No; 1=Yes Reference group=White
Investigator Race: Other	Race/Ethnicity of Investigators: Other	0=No; 1=Yes Reference group=White

Experience prior to DCP	Prior experience in DCFS before DCP initial start date (in month)	Continuous
Social Work Degree	Whether investigators have a Social Work degree	0=No; 1=Yes
Education level: Higher school or associate degree	Investigator's highest degree: Higher school or associate degree	0=No; 1=Yes Reference group=Bachelor's degree
Education level: Master's degree or higher	Investigator's highest degree: Master's degree or higher	0=No; 1=Yes Reference group=Bachelor's degree
Position: Children and family service intern	Investigator's position at DCP initial start date: Children and family service intern	0=No; 1=Yes Reference group=child protection specialist
Investigator caseload	Investigator's caseload during the investigation	continuous
Number of prior CPS reports	Total number of prior substantiated CPS reports of the family	continuous
Allegation: Sexual abuse	Allegation type in prior substantiated CPS report: Sexual abuse	0=No; 1=Yes
Allegation: physical abuse	Allegation type in prior substantiated CPS report: Physical abuse	0=No; 1=Yes
Allegation: Substance exposure	Allegation type in prior substantiated CPS report: Substance exposure	0=No; 1=Yes
Allegation: Emotional abuse	Allegation type in prior substantiated CPS report: Emotional abuse	0=No; 1=Yes
Allegation: Lack of supervision	Allegation type in prior substantiated CPS report: Lack of supervision	0=No; 1=Yes
Allegation: Environment neglect	Allegation type in prior substantiated CPS report: Environment neglect	0=No; 1=Yes

Allegation: Other neglect	Allegation type in prior substantiated CPS report: Other neglect	0=No; 1=Yes
Allegation: Substantial risk of harm	Allegation type in prior substantiated CPS report: Substantial risk of harm	0=No; 1=Yes