Apr 25th, 9:00 AM - 11:00 AM

Testing Assumptions about Laboratory Protocol Fidelity

Tess A. Simpson
East Tennessee State University

Wallace E. Dixon Jr.
East Tennessee State University

Albany Guyer
East Tennessee State University

Testing Assumptions About Laboratory Protocol Fidelity

Tess A. Simpson, Wallace E. Dixon, Jr., & Albany Guyer
*East Tennessee State University, Johnson City, TN 37614

Abstract

A recent focus of our Program for the Study of Infancy is the extent to which lab assistants retain fidelity when implementing experimental protocols. Dixon et al. showed that experimenters' implementations of some aspects of standardized protocols can be influenced by infants' temperaments. In this study, we evaluated archival videos involving experimenters' implementations of two elicited imitation tasks, Feed Bear and Make a Rattle. In both tasks ten experimenters were assumed to adhere to a narrative script and procedure. In this study we explored whether they did so. We examined two types of dependent measures reflecting 1) whether experimenters adhered to the familiarization time protocol, and 2) whether experimenters adhered to the standardized scripts. Results showcase the possibility that infant temperament may influence experimenter behavior and the importance of checking procedural fidelity.

Introduction

In human-subjects lab-based research, procedural fidelity can be compromised by the extent to which researchers unconsciously or inappropriately change their behavior based on their research participants. Procedural fidelity is a crucial element of behavioral lab-based research. Ellefson and Oppenheimer (2022) found that procedural deviations reduce effect sizes and lead to heterogeneity of findings over time. In developmental science, lab-based researchers typically assume that experimenters are not influenced by the unique characteristics of their research participants. At least, it is uncommon for researchers to publish procedural fidelity data. Dixon et al. (2021) found that experimenter' implementations of standardized protocols can be influenced by infants' temperaments, specifically looking–time. However, there is considerable literature on the effects of children on their social partners generally. Vallotton (2009) found that infants influenced their caregivers' responsiveness and quality of care through their own personal characteristics (i.e., gender, age, and communicative behaviors). Similarly, Snell et al. (2015) reported that toddlers who were perceived by their non-parental caregivers to have lower cognitive ability received less language stimulation. The purpose of the present study was to investigate whether developmental lab-based researchers were influenced in the laboratory by the children that participate in their research, specifically in deviations from the procedural script.

Method

Participants

Data were procured from an archival data set housed in the Program for the Study of Infancy at East Tennessee State University. In the data set, all participants were recruited from rural Southern Appalachia. Sixty-six primarily white, middle-class toddlers (M = 15.52 months, SD = 0.47 months) and their parents (61 mothers, 5 fathers) were recruited through local newspaper birth announcements. Approximately half of these participants returned 6 months later for a second lab visit (N = 32, M = 21.77 months, SD = 0.67), although these longitudinal data are not relevant for present purposes. Five participant videos were either (1) not recorded or (2) experimenters did not perform the task, resulting in 61 usable participant videos. These participants were comprised of 34 male and 27 female, primarily white toddlers (M = 14.86 months, SD = 2.06).

Materials

Infants' Temperament. The Early Childhood Behavior Questionnaire (ECBQ; Putnam, Garstein, & Rothbart, 2006) was used to assess temperament. The ECBQ measures temperament using parent-report and utilizes 201 items on an 8-point Likert-type scale (1 = never, 2 = very rarely, 3 = less than half the time, 4 = about half the time, 5 = more than half the time, 6 = almost always, 7 = always, 8 = not applicable). Caregivers rate the frequency of specific child behaviors over the previous two weeks. The items are then subsumed into 18 subdimensions. The subdimensions are further subsumed into three superdimensions: negative affectivity, surgency, and effortful control (Putnam, Garstein, & Rothbart, 2006).

Duration of Familiarization. Adherence to a target duration of object familiarization time (specifically, 60 seconds in each task) was evaluated as a function of infant temperament and linguistic proficiency. To extract this information, a team of two coders recorded familiarization time side by side in both the Make Rattle and Feed Bear tasks.

Infant Directed Speech (IDS) Talkativeness. We tracked experimenter IDS before the model (prologue IDS), during the model (narrative IDS) and after the model (epilogue IDS). In Feed Bear, the narrative script consisted of 98 words. In Make a Rattle, the narrative script consisted of 82 words. Transcribers used a two-pass procedure to reach reliability in transcribing IDS. Transcriptions were analyzed through Linguistic Inquiry and Word Count 2015 (LIWC) to evaluate the talkativeness (total words) of the IDS.

Results

Descriptive Statistics

Descriptive statistics for total words and familiarization time are reported in Table 1.

Table 1

Total Words and Duration of Familiarization by Feed Bear and Make a Rattle

	M	*SD*	*Min.*	*Max.*
Feed Bear				
TW				
Total	159.81	24.27	116.00	245.00
Prologue	11.45	8.94	0.00	35.00
Narrative	122.42	18.56	74.00	185.00
Epilogue	22.94	11.52	6.00	55.00
FAM	68.26	16.66	45.00	125.00
Make a Rattle				
TW				
Total	141.23	20.08	106.00	189.00
Prologue	12.63	8.56	0.00	36.00
Narrative	109.33	9.51	82.00	127.00
Epilogue	17.77	12.46	0.00	60.00
FAM	67.53	20.63	37.00	160.00

Note. TW = total words, FAM = Duration of Familiarization

Inferential Statistics

For familiarization times in the Feed Bear task, experimenters granted significantly longer familiarization times to the infants [M = 68.72, SD = 17.14; t(28) = 2.74, p = .011] than they were supposed to. But there were no differences among the experimenters. For Make a Rattle, there was no difference between the experimenter's familiarization times (M = 67.64, SD = 21.38) and expectations.

In both Feed Bear and Make a Rattle, experimenters used more words than they were supposed to [M = 123.90, SD = 18.03; t(28) = 7.73, p <.001] and [M = 109.00, SD = 9.52; t(27) = 15.01, p < .001] respectively. One-way ANOVAs revealed significant differences among experimenters in Make a Rattle, but not Feed Bear. There was a significant difference in word count by experimenter identity in Make a Rattle, F(6, 21) = 3.02, p = .027. There was no association between total IDS and infant temperament during the narrative phase.

Results (continued)

However, there were associations during the prologue and epilogue phases (See Table 2), especially for effortful control, sociability, and impulsivity.

Table 2

Correlations Between Temperament and Experimenter Total Words

	1	2	3	4	5	6	7
Both Tasks							
Total	-	-	-	-	-	.26	-
Prologue	.29	-	-	-	-	-	-
Narrative	-	-	-	-	-	-	-
Epilogue	-	-	.32	-.28	-	.23	-
Feed Bear							
Total	.36	-	.35	-	-	.36	-
Prologue	.47	-	-	-	.34	-	.40
Narrative	-	-	-	-	-	-	-
Epilogue	-	-	.44	-	.45	-	-
Make a Rattle							
Total	-	-	-	-	-	.33	-
Prologue	-	-	-	-	-	-	-.37
Narrative	-	-	-	-	-	-	-
Epilogue	-	-	-	-	-	-	-

Note. 1 = effortful control; 2 = positive anticipation; 3 = sociability; 4 = negative affectivity; 5 = activity; 6 = impulsivity; 7 = high intensity pleasure. Empty cells denote nonsignificant associations. Represented correlations were all significant at least the $p < .10$ level.

Conclusion

If infant characteristics, like temperament, are presumed to influence the quality of their social experiences, then it makes sense to hypothesize that experimenters would adapt their speaking and other behaviors to the child's characteristics (Snell et al., 2015; Vallotton, 2009). Dixon et al. 2021 also supported this hypothesis. The results reaffirm the possibility that infant temperament may influence experimenter behavior in the lab, in this case their IDS, although narrative IDS was not associated with temperament. This also reminds us of the importance of checking on procedural fidelity when implementing experimental protocols. Future research should continue to investigate the bidirectional influences of experimenters and their infant participants in terms of procedural fidelity.