Inconvenient samples: Modeling the effects of non-consent by coupling observational and experimental results

Yue Yu[12] (yue.yu@rutgers.edu), Elizabeth Bonawitz[2] (elizabeth.bonawitz@rutgers.edu),
Patrick Shafto[1] (patrick.shafto@gmail.com)
Department of [1]Math and Computer Science & [2]Psychology, Rutgers University - Newark
Newark, NJ 07102 USA

Abstract

Biased sampling of participants presents a major limiting factor for the generalizability of findings from behavioral studies. This effect may be especially pronounced in developmental studies, where parents serve as both the primary environmental input and decide whether their child participates in a study. To estimate the effects of parental non-consent, we coupled naturalistic observations of parent-child interactions with a behavioral test. Results showed that one particular parenting practice, the tendency to use questions to teach, associated with both children's behavior in the test and parents' tendencies to participate. Exploiting these associations with a model-based multiple imputation, we estimated that the means of the consented and not-consented groups could differ as much as 0.2 standard deviations for five of the seven test measurements we used, and standard deviations are likely underestimated. These results suggest that ignoring the role of consent may lead to systematic biases when generalizing beyond lab samples.

Keywords: sampling; generalization; parent-child interaction; learning; exploration; multiple imputation.

Introduction

Sampling and generalizability are the methodological bedrocks of science. Researchers often rely on measurements taken from a small group of volunteers to draw conclusions for a much broader population, so knowing whether the sample is representative of the population becomes critical to the validity and generalizability of research findings. Among the many factors that may bias the sampling process, one prevalent but under-studied factor is the refusal to participate in research. Because ethical treatment of research participants requires informed consent prior to their participation, we know very little about what characteristics are associated with non-consent, and what those who did not consent would have done if they had participated.

The problem is potentially more acute for fields in which behavior tends to be heterogeneous along factors that may associate with non-consent. One such field is experimental research with young children: On the one hand, before the start of schooling, children's experiences are heavily influenced by the values and practices of their parents, which are known to be heterogeneous both within and between social groups (Bornstein, 1991; Hoff, Laursen, Tardif, & Bornstein, 2002). At the same time, parents are also the ones who decide whether their children could participate in research, and the same values and practices may play a role in their decision. Given that parents' decisions are crucial for the composition of research

samples for most of the prevalent recruitment methods used in developmental experiments with young children (e.g., direct phone calls, recruitment from day care centers, preschools, and public spaces like museums), non-consent can present a major hurdle when evaluating the generalizability of findings from the field.

However, to date little is known about the factors associated with parents' non-consent to have their young children participate in experiments; consequently, it is difficult to speculate on the behavior of children who did not participate or on the implications for generalizability. This is in sharp contrast with the field of survey-based research with school-aged children and adolescence, where an extensive literature has associated parental consent and non-consent with both parents' characteristics and children's behavioral outcomes. For example, studies from 1970s-80s have shown that U.S. parents who are female, white and well-educated are more likely to return written consents for their children to participate in research (Kearney, Hopkins, Mauss, & Weisheit, 1983; Lueptow, Mueller, Hammes, & Master, 1977), and children who have better school performance and fewer behavioral problems are more likely to receive parental consents (Kearney et al., 1983; Severson & Ary, 1983). Moreover, the method of consent also matters. Compared to passive consent which requires a reply to opt out of a study, active consent which requires a reply to opt-in can bias the sample towards parents who are white and well-educated, and towards children who live in two-parent households, who have better school performance and satisfaction, who involve in more extracurricular activities and less risk-seeking behavior, and who are higher on self-esteem and assertiveness (Anderman et al., 1995; Dent et al., 1993). Critical to estimating these effects is the availability of relevant correlates, such as school records that contain demographic information and students' performance.

This study takes a first step to investigate whether parents' non-consent is also associated with preschoolers' behavior in standard experimental settings. Given that methods used to discover factors associated with non-consent in school-aged children (such as passive consent procedures and school records) are not usually applicable in the research with preschoolers, we developed a new method that is well-suited to many settings in which developmental psychologists collect data: coupling naturalistic observations of parent-child interactions with behavioral experiments.

By conducting the observation in public spaces without the awareness of the dyad, we aim to start with a relatively representative population that is unaffected by the consent

process. We then invite participation in a behavioral test to those who were observed. By analyzing the correlations between the observational and test data, and between the observational data and participation, we look for predictors that may associate test data from the participated group with participation itself, which would indicate the potential effects of non-consent on estimates of test measurements. If such association can be established, we would then use a model-based multiple imputation to simulate the behavior of children who did not participate, and compare the participated sample with the initial population on the test measurements.

The domain we chose to examine is one where there is known heterogeneity in parenting practices: the use of questions to teach. This line of research is grounded in a rich literature about informal pedagogy (Bonawitz et al., 2011; Csibra & Gergely, 2009), which suggests that the format in which parents and educators chose to present evidence to children influence how children infer and learn. Specifically, recent experiments (Yu, Bonawitz, & Shafto, under review) have shown that pedagogical questions asked by knowledgeable teachers are particularly effective in facilitating children's learning and exploration of a novel artifact. Given that the tendency to ask pedagogical questions has been shown to vary across parents (Yu, Bonawitz, & Shafto, 2016), we explore the effects of children's experiences with pedagogical questions on their responses in the experiment. We did so by replicating one condition of the previous experiment with an added observation phase, in which parents' pedagogical questions towards children were measured along with other parent-child interaction measurements. This allows us to look for associations between parents' pedagogical questions and children's responses to these questions. And because the observational data is available for children who did not participate, these associations could then be used to estimate the test data for the whole population.

Method

Testing sites

We set up the study in two sites: an indoor reptile exhibit in a zoo, and an indoor playground. The zoo is in Essex County, NJ, which is one of the nation's most racially diverse counties, and has one of the nation's most unequal economies measured by the Gini Index (U.S. Census Bureau, 2016). The playground is in Middlesex County, NJ, which is overall more affluent, but also has a highly racially diverse population, of which 30% were born outside the U.S. (U.S. Census Bureau, 2016). We chose these two sites to ensure diversity in the population we initially observed. Consents from these two sites and from the internal review board was obtained before conducting the study.

These two sites differ in the expected level of supervision and involvement from parents. The zoo is an open environment that requires parents to constantly supervise their young children. Exhibits in the zoo also feature many textual materials which require parents' explanations for young children to understand. In comparison, the playground features spaces and activities which young children can navigate on their own, and the closed environment allows minimal supervision from their parents. This contrast allows us to test whether the characteristics associated with parental non-consent differ by the type of facilities from which they were recruited.

Participants

Between the two sites, we observed a total of 109 parent-child dyads. Of these 109 dyads, 31 were not invited for the test because of one of the following: the dyad left before the observation was finished (18), parent interrupted researchers during observation (1), researchers did not get a chance to invite parent (4), adult accompanying child was not the child's parent (4), child was out of the target age range (2), or child did not speak English (2). The remaining 78 parent-child dyads comprised our "population", which is unaffected by the consent process. Among them 41 were recruited from the zoo, and 37 were recruited from the playground.

Procedure

During each trip to the testing sites, three researchers collected data from parent-child dyads in three phases: Two coders first observed and coded the interactions between the parent and the child (observation phase). Then a third researcher invited the dyad to a test (recruitment phase). She and one of the coders conducted the test if the dyad agreed to participate (testing phase).

Observation phase. Two coders pretended to be visitors of the zoo or the playground, so that they could code parent-child interactions without the dyad's awareness. The coders first looked for a child who was estimated to be between 3 and 6 years of age, and determined who were the adults accompanying the child. If at least one adult looked like the child's parent, the coders would record the members in the group (e.g., father, mother, and two daughters), and agree on a target dyad for observation (e.g., mother and younger daughter). To reduce potential selection biases from the coders, they always observed the first dyad they saw that fitted the requirements, and the observation always started immediately once the target dyad was determined. Each dyad was observed for 5 minutes, during which the coders independently coded both the quantity and the quality of parent-child interactions. Quantity of interaction was measured by the length of time period of dyadic activities (parent and child engaging in the same activity), supervised activities (parent watching, following, or taking pictures of child while child is engaging in his or her own activities), and unsupervised activities (parent and child engaging in different activities). We coded these as mutually exclusive categories, and they added up to the total time length (5 minutes). Quality of interaction was coded as a set of frequency measurements, adapted from the Dyadic Parent–Child Interaction Coding System (Eyberg, Nelson, Ginn,

Bhuiyan, & Boggs, 2013): The coders recorded the numbers of parents' questions, statements, and commands towards children.

Critical to our interest, parents' questions were further differentiated based on their functions (Yu et al., 2016): Those used to help children learn were coded as "pedagogical questions", whereas those used to request information from children were coded as "information-seeking questions". All coders were trained for approximately 5 hours and practiced the coding scheme with at least 5 parent-child dyads before formal data collection. Inter-rater reliabilities were high for both the quantity and quality of parent-child interactions: Inter-rater correlation $r = .78 \sim .84$ for quantity codes, and $r = .79 \sim .86$ for quality codes. The average of the two coders' codes were used for data analysis.

Recruitment phase. After the 5-minute observation, a third researcher who was blind to the observation phase approached the parent and invited the parent-child dyad to participate in a test. The recruitment procedure followed a script which resembled that of a typical developmental experiment: The research started with a brief self-introduction, then described the research as a study of how children learn and explore a novel toy, then briefly explained the consent form, and finally asked if the parent would be interested to have his or her child participate in the test. For parents who had multiple children with them, we specifically asked for the child who had been observed. Among the population of 78 parent-child dyads that were observed, 59 agreed to participate (the "consented" group) and 19 refused (the "not-consented" group). Of the 59 parents who agreed, 11 children did not participate in the test (2 were busy playing and did not get a chance to come, 8 refused to come, and 1 did not understand English), and the video was missing for one additional child, so data from the testing phase were available for 47 children. According to parental report, children who participated in the testing phase were diverse regarding race (51% white, 4% black, 15% Hispanic-Latino, 13% Asian, 17% multi-racial), but most came from middle- or upper-class families (91% of the main caregivers have college diploma or above, 84% of the families have annual house hold income of $50K or above).

Testing phase. Parents and children who agreed to participate were led to a corner of the zoo exhibit or a separate room in the indoor playground, where the test was conducted by the recruiter (acting as an experimenter) and one of the coders (acting as a confederate). The materials and procedure of the test was adapted from Bonawitz, et al. (2011), and was identical to the pedagogical question condition in our recent experiment (Yu et al., under review).

A novel toy of approximately 14" × 7.5" × 14.5" was used in the test. In addition to several inert properties, the toy had five functional parts: a tower that lit up when a button was pushed, a knob that produced a squeaking sound when squeezed, a lady bug pin light that flashed in three different patterns when pushed, a flower magnet that moved between three different places on the toy, and a turtle hidden in a pipe that was visible through a magnifying window.

During the test, the child sat at a table opposite the experimenter and the confederate. The toy was initially hidden out of sight. The experimenter first said that she knows about the toy and the confederate does not, and asked the confederate to bring out the toy. After the confederate brought out the toy and handed it over to the her, the experimenter then asked a pedagogical question to the child, "I'm asking you to think about: What does this button do?", while pointing to the button on the tower without activating it. Then she told the child it is his or her turn to play with the toy, and to let the researchers know when he or she is done. The test ended when the child stopped playing and signaled the researchers, and a sticker was presented as a reward. The whole phase was video recorded.

Video coding

After data collection, the videos from the testing phase were coded by another research assistant who was blind to the observation phase and to the hypotheses of the study. She first determined the total time children spent playing with the toy, and then coded three measurements regarding both the whole playing period, and the first minute after children started playing: whether children activated the target function (the tower with the button), the number of unique actions they performed with the toy, and the number of non-target functions (out of 4) they activated. A second coder coded 14 (30%) of the videos, and the inter-coder reliability was high for all measurements: total time playing: $r = .98$; activating target function: Cohen's $\kappa = 1$ for both total time and first minute; number of non-target functions activated: Cohen's $\kappa = 81$ (total time) and $\kappa = 75$ (first minute); number of unique actions performed: $r = .79$ (total time) and $r = .92$ (first minute).

Results

Our population consisted 32 mother-son dyads, 16 mother-daughter dyads, 17 father-son dyads, and 13 father-daughter dyads. Parent-child interactions varied both across sites and within sites: Compared to dyads in the playground, dyads in the zoo spent more time on dyadic activities, and less time on supervised (but not dyadic) activities or unsupervised activities, ts > 2.6, ps $< .01$. Parents also asked more pedagogical and information seeking questions, and said more statements in the zoo than in the playground, ts > 3.4, ps $< .001$. The difference in parents' commands toward children was marginally significant, $t(67.7) = 1.75$, $p = .09$. These results suggest that the testing site needs to be taken into account when interpreting parent-child interactions. Therefore, testing site had been entered as a control variable for all further analyses. We also observed large within-site variations: For all measurements, standard deviations were higher than 1/3 of the mean for both the zoo and the playground. These variations suggest that the population we observed was diverse with regard to parent-child
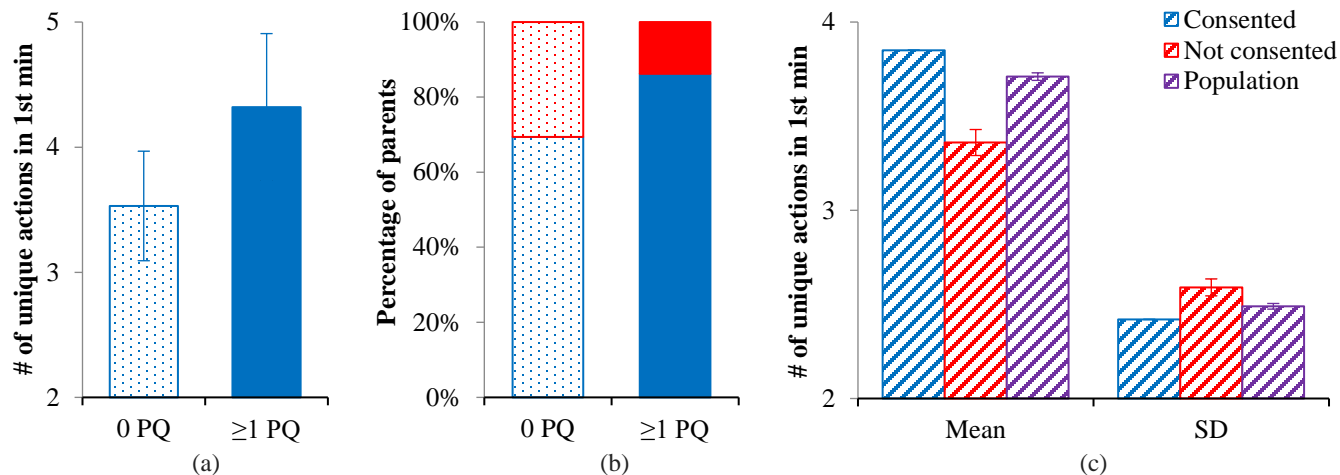
Figure 1. The estimated effects of non-consent on one of the test measurements: the number of unique actions children performed during the first minute of play. We tested whether the mean and the standard deviation measured from the consented group (1c, blue bars) are unbiased estimates of the population. To do so, we separated the consented group according to the number of PQs parents asked children during the observation prior to the test (1a). Parents who asked more PQs tended to have children who performed more unique actions during the test. On the other hand, as shown in (1b), parents who asked more PQs also tended to consent their children to participate (blue) rather than not consent (red). These two associations resulted in the simulated not-consented group to have a lower mean than the consented group (1c), by an estimated effect size (Cohen's d) of 0.2. Compared to the population (purple bars), focusing on the consented group may result in an overestimation of the mean and an underestimation of the standard deviation. PQ = pedagogical questions. Error bars denote standard errors across children (1a) or across simulations (1c).

interactions, which serves as a basis for further correlational analyses. Though parent-child interactions differed by site, the proportion of parents who agreed to participate did not differ significantly, playground: 25 agreed, 12 refused (68% vs. 32%); zoo: 34 agreed, 7 refused (83% vs. 17%); Fisher's exact $p = .19$.

Are parent-children interactions associated with children's behavior in the test?

Test data was available for 47 children ranging from 3.0y to 6.3y, of which 27 were recruited from the zoo and 20 from the playground. Children from the two sites did not differ with regard to the activation of target and other functions, or the number of unique actions they performed on the toy, $ts < 1.4$, $ps > .1$. However, there was a trend of children playing longer with the toy in the playground than in the zoo, $M_{zoo} = 189s$, $M_{playground} = 132s$, $t(29.4) = 1.83$, $p = .08$, $d = 0.58$. When comparing these results with previous experiments we conducted in preschools ($n = 30$, age range = 4.0y to 6.0y) using the exact same protocol (Yu et al., under review), none of children's response measurements differed significantly across the three sites, $Fs < 2.2$, $ps > .1$.

Next we looked at the relation between children's responses during the test and parent-child interaction measurements during the observation. After controlling for testing site and age, measurements regarding the composition of the group being observed (parent' and child's gender, and whether they were accompanied by other adults or children) did not correlate with any of children's responses, $ps > .1$. However, measurements of parent-child interaction did correlate with children's responses: Children of parents who spent more time

watching and following them were less likely to discover the target function during the first minute of play, $r(42) = .33$, $p = .02$. At the same time, children whose parents asked more pedagogical questions discovered more other functions of the toy, $r(42) = .32$, $p = .03$, and also performed more unique actions during first minute of play (Figure 1a), $r(42) = .29$, $p = .05$. These results suggest that patterns observed in parent-child interactions were indeed associated with children's learning and exploration during the test.

Are parent-child interactions associated with participation?

We then examined whether patterns observed in parent-child interactions also predicted parents' responses to the invitation for research. We fitted a logistic regression model with participation as the dependent variable and the observational measurements as the predictors. Overall the model predicted actual participation with 80% accuracy. With regard to individual predictors, parents were more likely to have their boys participate than girls, $B = 1.47$, $p = .03$; and those parents who asked more pedagogical questions during the observation were more likely to participate, $B = 1.49$, $p = .05$ (Figure 1b).

What can be predicted for children who did not participate?

Results so far have shown that the number of pedagogical questions parents asked children predicted both children's participation in a test and their behavior during the test. This indicates that children's participation and behavior may be related as well—that is, if we have tested children whose

parents did not consent them to participate, they may have responded differently than children who did participate.

To test this hypothesis, we applied model-based multiple imputation to our data (Rubin, 2004).[1] The model we used for multiple imputation was a stochastic regression model, implemented with IBM SPSS 22. The seven observational measurements were used to model the seven test measurements, based on data from the consented group. The resulting models were then used to predict behavior of the not-consented group stochastically (with random noise) for 100 independent runs of simulations.[2]

Results showed that across the 100 runs of simulations, the means of the not-consented group were consistently different from that of the consented group for five out of seven test measurements including activating target function (total time and first minute), number of non-target functions activated (first minute), and number of unique actions performed (total time and first minute). The departure was towards the same direction—the participated children learned and explored more with the toy (Figure 1c shows one example). The differences between the means of the consented and not-consented group were estimated to be between 0.09 and 0.20 standard deviations for these five measurements. In addition, compared to the population, focusing on the consented group alone would lead to consistent underestimation of the standard deviations across children, and this is true for all test measurements we examined.

Discussion

This study takes a first step towards evaluating whether

results from children who participated in an experiment could generalize to children whose parents did not consent for them to participate. We attempted to estimate these potential biases with a novel approach by pairing a behavioral test with naturalistic observations of parent-child interactions prior to parental consent. Results have shown that a specific parenting practice—asking questions to help children learn—correlated with both parents' tendencies to have their children participate in the test, and children's learning and exploration during the test. And since the observational data was available for both those who participated and did not participate, we were able to exploit these associations to impute behavior for children who did not participate. Results from the imputation showed differences in group means between the consented and not-consented group for five out of the seven test measurements, with estimated effect sizes (Cohen's ds) between 0.09 and 0.20. Furthermore, the consented group showed a lower standard deviation than the population for all test measurements.

Before discussing the implications of the results, it is worth noting that several assumptions underlie these simulated estimates. First, we assumed no direct causal relation between parents' decisions to have their children participate and children's potential behavior in the test. This assumption is plausible in our case: Because parents were not given much detail about the testing procedures, their decision to participate is unlikely to be based on what they expect their children to do. However, in other situations this assumption could be violated, which could render the imputation analysis invalid. For example, in a study that measures children's executive functions, if children drop out from the study exactly because of low executive functions, then it would be invalid to impute executive functions for the dropout group even when all relevant correlates have been observed and entered into the model. Second, our approach is valid because we saw variations in parent-child interactions for both the consented and not-consented groups, as well as significant overlap between the two groups. This allows imputation to be done as interpolations within the ranges of empirical support. In cases where the consented and not-consented groups do not overlap, our approach could be invalid, as the relations found in the consented group may not extend to the not-consented group. In sum, our methods to generalize experimental results are themselves subject to usual conditions for generalization.

How much this new approach could be and should be implemented in developmental experiments would also depend on various factors. The first factor is the recruitment method. Our approach could be beneficial for research settings that provide opportunities to observe and recruit from a relatively diverse population, such as in public spaces. On the other hand, for studies recruiting from places with a preselected population, such as preschools, the demographics of the preselected population may present a stronger sampling bias than parents' consent. The second factor is the research topic. Our approach could be more

---

[1] Multiple imputation is the recommended tool to predict missing data when missingness depended on other observed variables, but not the missing variable itself (Sinharay, Stern, & Russell, 2001). In our case, because missingness was a result of the parent's decision, it was associated with patterns observed in parent-child interactions (as shown in the logistic regression), but was not directly associated with children's behavior in the test. Therefore, multiple imputation is suitable to simulate behavior for the not-consented children.

[2] The two test measurements regarding activation of target function were imputed as binary variables, whereas all other test measurements were imputed as continuous variables. For each test measurement, a logistic regression model (for binary variables) or a linear regression model (for continuous variables) was first fitted on the data from the consented group. From the fitted model, posterior distributions were computed for the 8 parameters in the logistic regression model or the 9 parameters in the linear regression model (intercept, coefficients for the 7 observational measurements, and the residual variance for linear regression only). Then the values for the not-consented group were imputed for $m = 100$ runs. For each run, a new set of parameters were randomly drawn from their respective posterior distributions, and were used to compute the expected values plus random errors for each child in the not-consented group. The means and standard deviations of the not-consented group and of the whole population were then calculated for this test measurement. The procedure is then repeated for the remaining runs of simulations, and applied to the other test measurements.

valuable for domains in which parent-child interactions have been, or are expected to be, associated with children's behavior. The last factor is the research ethics. Pre-consent observations are ethically viable only for public actions, and needs to be performed with caution.

In cases where our approach can be applied, it could benefit the interpretation and generalization of experimental findings in several ways: First, it could reveal correlations between parent-child interactions and children's behavior, which may help explain the cognitive mechanisms and environmental inputs associated with the observed behavior. Second, it could inform the generalizability of experimental findings to children whose parents did not consent them to participate. Third, it can serve as an empirical base for future research to recruit a more representative sample. By knowing the associations between parental consent and patterns in parent-child interaction, it may be possible to intentionally focus recruitment on parent-child dyads who are likely underrepresented in typical recruitment procedures.

Our results may also have implications for developmental theories. Many developmental theories are built upon findings from experiments, as experimental design has advantages in addressing a range of developmental questions: These include depicting developmental trajectories ("Children do X at age Y"), disentangling causal mechanisms underlying children's behavior ("Children do X because of Z"), and testing causal effects of interventions ("T helps children do X"). In typical cases, random assignment of participants across groups removes unwanted systematic differences between groups, so that the effects of age, condition, or treatment can be detected by comparing between-group differences with within-group differences. Our results have shown that parental non-consent may have biased this comparison in two ways that random assignment cannot solve: First, it could lead to an underestimation of within-group variations, and thus Type I errors may be underestimated and effect sizes may be overestimated. Second, compared to the general population, children who received consent may be more susceptible or insusceptible to certain manipulations or treatments, therefore biasing the estimation of the between-group differences. Because findings from developmental experiments often guide real-world practices which apply to the general population, understanding factors and biases associated with non-consent is essential when interpreting and applying these findings.

To conclude, this study provided a first empirical demonstration that children with and without parental consent to participate in research may have differed in behavior measured in an experiment. Therefore, parental non-consent should be considered an important factor when evaluating the generalizability of experimental findings, and the theories built upon them. In addition, we provided a method that, in certain contexts, could be used to estimate the effect of parental non-consent and generalizability of experimental results.

References

Anderman, C., Cheadle, A., Curry, S., Diehr, P., Shultz, L., & Wagner, E. (1995). Selection bias related to parental consent in school-based survey research. *Evaluation Review, 19*(6), 663-674.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition, 120*(3), 322-330.

Bornstein, M. H. (1991). *Cultural approaches to parenting*. Hillsdale, NJ: Lawrence Erlbaum.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148-153.

Dent, C. W., Galaif, J., Sussman, S., Stacy, A., Burtun, D., & Flay, B. R. (1993). Demographic, psychosocial and behavioral differences in samples of actively and passively consented adolescents. *Addictive Behaviors, 18*(1), 51-56.

Eyberg, S., Nelson, M., Ginn, N., Bhuiyan, N., & Boggs, S. (2013). Dyadic parent-child interaction coding system (DPICS): Comprehensive manual for research and training (4th ed.). Gainesville, FL: PCIT International.

Hoff, E., Laursen, B., Tardif, T., & Bornstein, M. (2002). Socioeconomic status and parenting. *Handbook of parenting Volume 2: Biology and ecology of parenting, 8*(2), 231-252.

Kearney, K. A., Hopkins, R. H., Mauss, A. L., & Weisheit, R. A. (1983). Sample bias resulting from a requirement for written parental consent. *Public Opinion Quarterly, 47*(1), 96-102.

Lueptow, L., Mueller, S. A., Hammes, R. R., & Master, L. S. (1977). The impact of informed consent regulations on response rate and response bias. *Sociological methods & research, 6*(2), 183-204.

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys*. Hoboken, NJ: John Wiley & Sons.

Severson, H. H., & Ary, D. V. (1983). Sampling bias due to consent procedures with adolescents. *Addictive Behaviors, 8*(4), 433-437.

Sinharay, S., Stern, H. S., & Russell, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological Methods, 6*(4), 317.

U.S. Census Bureau. (2016). 2011-2015 American community survey. Retrieved from factfinder.census.gov

Yu, Y., Bonawitz, E., & Shafto, P. (in press). Pedagogical questions in parent-child conversations. *Child Development*.

Yu, Y., Bonawitz, E., & Shafto, P. (under review). Questioning supports transmission of knowledge and exploratory learning in pre-kindergarden children.