HHS Public Access

Author manuscript

Appl Dev Sci. Author manuscript; available in PMC 2021 January 01.

Published in final edited form as:

Appl Dev Sci. 2020 ; 24(3): 242–262. doi:10.1080/10888691.2018.1462090.

Retrospective Report Revisited: Long-Term Recall in European American Mothers Moderated by Developmental Domain, Child Age, Person, and Metric of Agreement

Marc H. Bornstein, **Diane L. Putnick**, **Kyrsten M. Costlow**, **Joan T. D. Suwalsky**

Child and Family Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Public Health Service

Abstract

Retrospective parental reports are common in the developmental science literature, but their validity has been questioned. We investigated the consistency of retrospective maternal recall by comparing original with retrospective maternal reports in three domains (maternal cognitions, mother-reported child and mother behaviors, and observed child and mother behaviors) at three retention intervals (12, 14, and 15 years) in two metrics (individual standing and group level). In a longitudinal study, European American mothers (*N*=46) provided data when their children were 5, 20, and 48 months of age and retrospective recall data for each age when their children were 16 years. Overall, mothers recalled similar average mean levels (49% of variables explored) or better mean levels (41% of variables) retrospectively; better levels indicating a positive recollection bias. At least moderate consistency in relative standing was evident for 52% of variables. Still, the findings varied somewhat by domain, child age, and person. Retrospective parental reports can provide accurate accounts of the past, but should be used with caution, as their consistency varies and is specific to moderating factors.

Keywords

maternal report; retrospective; recall; validity; child development

Parental report is a common method for assessing child development and parenting and can take the form of interviews, questionnaires, checklists, or focus groups (Bornstein, 2014). Some developmental literatures – child language, adaptive behavior, and temperament, for example – rely heavily on parental report (Feldman et al., 2005; Fenson et al., 1994; Keenan & Wakschlag, 2002; Keenan et al., 2007; Sparrow, Balla, & Cicchetti, 1984). Indeed, parental reports are sometimes preferred to testing and observational methods in developmental science for a variety of reasons. Parental reports are rapid, efficient, and cost-effective. With respect to child development, parental reports gather information from those closest to the child and so are thought to draw on a wealth of knowledge unattainable through testing or observation (Bates, Bretherton, & Snyder, 1988; Bornstein & Toole, 2010;

Address correspondence to: Dr. Marc H. Bornstein, Child and Family Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, Suite 220 6555 Rock Spring Drive, Bethesda MD 20817 USA, TEL: 301-496-6832 FAX: 301-496-2766, Marc_H_Bornstein@nih.gov WWW: http://www.cfr.nichd.nih.gov.

Thal & Bates, 1990). Parental reports are particularly useful in early development when certain behaviors are infrequent or difficult to test or observe (Bates & Carnevale, 1993; Bowerman, 1985) and with children who struggle to cooperate in a testing environment (Feldman et al., 2005). Furthermore, for studies of parenting practices, only parent self-reports can provide crucial information on personal and subjective experiences as parents. Parental reports of parents' perceptions are also valuable because they constitute the climate surrounding the child (Bornstein, Gaughran, & Seguí, 1991). Parents' recollections of their children's development and their own parenting may color their current conceptions of both. For these several reasons, knowing more about the validity of parents' reports is valuable.

All that said, parental reports are sometimes thought to be less valid (in the criterion-related validity sense) than testing and observation. Because parents are untrained, their reports may be unsystematic or unreliable. Although possessing a wealth of knowledge about their children's development, parents (like other reporters) are also possibly biased (Bornstein & Toole, 2010; Feldman, Wentzel, & Gehring, 1989). For example, social desirability can encourage parents to distort their responses and cast their children in a more positive light (Kagan, 1998). Parents (especially first-time parents) may also lack the experience required to accurately judge their children's development in relation to other children (Forman et al., 2003; Seifer, Sameroff, Barrett, & Krafchuk, 1994). Furthermore, parental reports might be influenced by a variety of other factors, such as parent age, intelligence, education, SES, and personality (Bates & Bayles, 1984; De Los Reyes & Kazdin, 2005; Sameroff, Seifer, & Elias, 1982; Vaughn et al., 1987). Child characteristics, such as age, gender, health status, birth order, personality, and social competence, may also shape parent report accuracy (Bornstein, 2014).

Retrospective Reports

One common variant of parental report is the retrospective report. Retrospection has been employed for a wide variety of memoranda, although controversy exists around the reliability and validity of information derived from retrospective reports as well (Widom & Morris, 1997). Classical and contemporary authorities in memory research have often cautioned against testimony based on recall because memory is partly a reconstructive (as opposed to reproductive) system (Allport & Postman, 1947; Bartlett, 1932; Ebbinghaus, 1913; Stern, 1902) and may be influenced by later information and by an individual's attitude, mood, and mental state (Bower, 1981; Lewinsohn & Rosenbaum, 1987; Matt, Vázquez, & Campbell, 1992; McFarland & Buehler, 1998; Rutter, Maughan, Pickles, & Simonoff, 1998). Some issues in establishing accuracy of retrospective recall include the sheer fallibility of memory (Brennen et al., 2010; Widom & Morris, 1997; Williams, 1994) and measurement bias (Fergusson et al., 2000). Recall can be distorted at different stages of processing, including initial observation, encoding, retention, and communication (Yarrow, Campbell, & Burton, 1970).

The present study concerns mothers' retrospective recall of diverse child and parent factors in their children's early development. In the seminal paper related to this topic, Yarrow et al. (1970, p. 68) compared retrospective reports with childhood records, found that errors in recall reflected mothers' current perspectives on their children's characteristics and

personalities, and concluded that their analyses “brought little honor to retrospective reports as records of children’s development.” Indeed, the rather poor contemporary reputation of retrospective parent reports springs largely from this summary observation. A more recent study looked at parent recall in the health domain. Bone fractures are common in childhood; approximately one in three children will sustain at least one fracture by 18 years of age (Cooper, Dennison, Leufkens, Bishop, & van Staa, 2004; Jones, Williams, Dow, & Goulding, 2002). The validity of parentally reported lifetime fracture prevalence in children is important to determine because the accuracy of parental report is vital to making a diagnosis of osteoporosis in children. Moon et al. (2016) found that approximately one in six injuries that parents reported to be a fracture was not a radiographically confirmed bony injury, and that one in five parental fracture histories contained incorrect information on either fracture occurrence or fracture site, or fractures were forgotten.

Despite this array of criticisms, several studies have reported that retrospection can be accurate (Hardt & Rutter, 2004; Paivio, 2001). After all, five in six (83%) fractures that parents reported were confirmed radiographically, and four in five (80%) parental fracture histories contained correct information on fracture occurrence and site (Moon et al., 2016). Indeed, according to Brewin, Andrews, and Gotlib (1993), many claims have exaggerated the unreliability of retrospective reports. Some studies have shown retrospective reports to be both reliable and valid, suggesting that they may not be as psychometrically poor as often billed (Bernstein et al., 1994). In the population-representative Dunedin cohort (N = 1,037), Reuben et al. (2016) estimated agreement between adverse childhood experiences prospectively recorded throughout childhood by staff at ages 3, 5, 7, 9, 11, 13, and 15 years and retrospectively recalled in adulthood by study members when they reached age 38. Retrospective and prospective measures of adversity showed moderate agreement.

Retrospective Reports of Parenting and Child Development

Child development embraces multiple domains of health, motor, cognitive, and socioemotional growth just as parenting is instantiated in diverse cognitions and practices. In consequence, both child development and parenting are multidimensional, modular, and specific (Bornstein, 2015). In considering mothers’ recall of their children’s development or their own parenting, therefore, it is unlikely that retrospection would be identical for all domains and for all cognitions and practices. Rather, recall more likely varies with the child and parent topic. This circumstance underscores the values of identification and empirical focus on different domains of child development as well as individual cognitions and practices of parenting. It could be, for example, that some developmental domains in children are recollected better than others just as some parenting cognitions and practices in development are recollected better than others. Without testing multiple variants of each (as we do here), we simply do not know and, worse, risk tarring all with the same broad brush.

In this study, we focused on mothers’ explicit or declarative memory, one main type of long-term human memory that is the conscious, intentional recollection of factual information from previous personal experiences. In specific, we assessed episodic (as opposed to semantic) memory, which stores specific personal autobiographical experiences that is memories for events which one has personally experienced (the contextual who, what, when,

where, why) that can be explicitly stated (Tulving, 2005). The memories we were interested in all collect past personal experiences that occurred at particular times and places. Recollection is a main component of episodic memory, retrieval of information pertaining to specific life events or experiences. Such autobiographical memories for personal events are encoded in everyday life, are elicited with word or event cues, and may be retrieved after intervals of years or decades (Conway and Pleydell-Pearce, 2000; Pillemer, 2003; Rubin, 2006). Typically, autobiographical memory research distinguishes among components of events and measures lengthy intervals (Conway, 2009; Eichenbaum et al., 2001; Piolino et al., 2009).

The retrospective memories we sought to tap in mothers concerned their children's early life and development and their own parenting responsibilities. Retrospective reports are likely influenced by the emotional intensity of the events being reported (Barsky, 2002), and parenting is unquestionably an emotional bit of business (Dix, 1991; Leerkes & Augustine, in press). Emotions can improve episodic memory (Holland & Kensinger, 2013), from initial encoding to consolidation to retrieval (Dolcos et al., 2012; LaBar & Cabeza, 2006). Emotions increase the likelihood that an event will be remembered later and that it will be remembered more vividly. Memories are also enhanced because of the propensity to rehearse and retell emotional events (McCloskey, Wible, & Cohen, 1988).

Of course, recall of information is influenced by the way in which to-be-remembered information is processed. "Depth-of-processing" refers to improvement in recall about information a person has pondered or has meaning for the person. Processing information with increased emotional or social or personal relevance engages more effortful elaboration (Sakaki, Niki, & Mather, 2012), and information with emotional or social or personal relevance enhances memory for the information encoded with reference to oneself (called the "self-reference effect;" Macrae et al., 2004; Yaoi, Osaka, & Osaka, 2015). Emotional memories are often more detailed and stronger than neutral ones, and memories for emotional experiences are also known to increase with time, whereas neutral memories may be promptly forgotten (Mather & Sutherland, 2011; Yonelinas & Ritchey, 2015).

This Study

With all this in mind, here we re-visit the question of maternal retrospective recall by examining the consistency of retrospective parental reports with mothers' original reports obtained in the context of a prospective longitudinal study. As retrospective parent reports have been widely used and may be helpful to the study of child development and parenting, the criterion validity of reports has important implications for developmental and clinical research. The present study seeks to address past conflicting viewpoints about retrospective reports by re-investigating consistency between retrospective maternal reports and mothers' original reports in a small number of novel ways:

1. 3 Domains. To build on and differentiate the extant literature, we collected maternal reports covering three domains: maternal cognitions, mother-reported child and mother behaviors, and observed child and mother behaviors.

2. 3 Times. Reports from memory may be influenced by the amount of time between the event and recall. We compared information from mothers collected when their infants, toddlers, and children were 5, 20, and 48 months, respectively, to retrospective reports about the same information collected at 16 years during their children's adolescence.

3. 2 Metrics. We focused on two measures of consistency: group mean levels of mothers' reports between times (*t*) and individual relative standing of mothers' reports between times (*r*). The two measures of consistency are conceptually and statistically independent (Bornstein, Putnick, & Esposito, 2017). Group mean level consistency measures whether there are no systematic differences in recollections in the group as a whole (versus a consistent positive recollection bias, for example). Relative-standing consistency (i.e., correlation) measures similarity in mothers' original and retrospective reports in their ordering within the group at both measurements.

4. 3 Moderators. We examined potential explanatory variables of three sorts: maternal, child, and situational characteristics.

5. Confidence. We collected mothers' confidence ratings of their retrospective reports.

We focused on maternal report specifically because mothers have traditionally and cross-culturally assumed primary responsibility for childrearing (Bornstein, 2015; Murray, Richards, & Nihouarn-Sigurdardottir, in press; Parke & Cookston, in press). This is the only study to our knowledge to examine the consistency of retrospective reports on different child developmental domains as well as different parental cognitions and parenting practices spanning so long periods and from different statistical vantages.

Method

Participants

As part of a prospective, longitudinal study, 46 European American mothers provided data when their firstborn children (17 girls, 29 boys) were infants (M = 5.38 months of age, SD = .18; range = 4.86 – 5.82), toddlers (M = 20.09 months of age, SD = .20; range = 19.71 – 20.67), and preschoolers (M = 48.24 months of age, SD = .87; range = 46.39 – 49.91). Then, when their children were adolescents (M = 16.09 years of age, SD = .60, range = 15.15 – 17.23 years), mothers provided retrospective data about their children and their parenting when their children were 5, 20, and 48 months. At the time of retrospective data collection, mothers averaged 48.08 years of age (SD = 4.10, range = 41 to 58 years). All but 8 mothers worked at least part time, and 73% had completed a 4-year college degree. Family socioeconomic status averaged 52.59 (SD = 9.15) on the Hollingshead (1975) scale. At the beginning of the study (when children were 5 months), all mothers were married. By the time of retrospective data collection, 23% of mothers were separated, divorced, or remarried. Families also varied in number of children: 30% had only the one child, 47% had two children, and 23% had three or more children.

Procedures

Institutional Review Board approval was obtained from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, protocol 88-CH-0032, titled, "The influence of maternal age, employment status, and parenthood status on children's cognitive development." Due to funding and concerns about participant burden, 70 mothers from this larger longitudinal study of families with firstborn children ($N = 317$) were approached to participate in the retrospective portion of the study. The longitudinal study was designed to track normative child and family development. This retrospective study was added between two longitudinal waves at 14 and 18 years to take advantage of the longitudinal design. One mother refused. Those who responded and agreed to participate ($n = 69$; 98.57%) were mailed a questionnaire asking them to recollect information about themselves, their children, the birth of the child, support networks for parenting, history of maternal employment, and associated substitute childcare that they had provided when their children were 5, 20, and 48 months. Most mothers ($n = 46$; 66.67%) returned the questionnaire by mail. Family socioeconomic status and maternal education were similar in the final sample of 46 mothers ($M = 51.17$, $SD = 8.80$ and $M = 5.85$, $SD = .87$, respectively) and the remainder of the original sample ($n = 271$; $M = 51.87$, $SD = 13.55$ and $M = 5.64$, $SD = 1.44$, respectively), $t(86.25) = -.45$, $p = .652$ and $t(93.67) = 1.35$, $p = .180$, respectively. However, mothers in the final sample were older when their children were 5 months ($M = 32.45$, $SD = 4.32$)than mothers who did not participate ($M = 29.45$, $SD = 6.56$), $t(85.09) = 3.98$, $p < .001$.

The questionnaire was divided into five parts: current demographics, questions about mothers than mothers who did not participate ($M = 29.45$, $SD = 6.56$), $t(85.09) = 3.98$, $p < .001$.and children at 5, 20, and 48 months, maternal employment, childcare, and social support. Mothers were also asked to provide confidence ratings for each part. The original data included completing questionnaires and participating in multiple videorecorded laboratory and home visits and interviews. The retrospective survey contained selected measures from the original data collection points designed to assess maternal recall of (1) her own cognitions, (2) her reports of her child's and her own behavior, and (3) observed child and mother behavior when her child was 5, 20, and 48 months of age. When possible, questions in the retrospective survey were asked identically to the way they were asked originally. However, some original assessments were too long and/or measures were taken via behavioral observation or experimenter assessment and were therefore adapted for the retrospective survey. In each section of the questionnaire, the retrospective questions were put in context by asking mothers, for example, to "think about your baby when s/he was 5 months old" or "think about your style as a mother when your child was 20 months old." Where appropriate, internal consistency of scales (Cronbach's alpha) is reported in Tables 1 and 2. Additional psychometric reliability and validity data on all measures are available from the first author.

Domain 1 Measures: Maternal Cognitions

Mothers' memories of their cognitions were composed of measures that were re-administered either verbatim or with only minor wording modifications. Two separate questions assessed maternal memories of *satisfaction with the parenthood-employment*

balance and *satisfaction with childcare arrangements* when children were 5 and 48 months. The response format was 1 = *very dissatisfied*, 2 = *dissatisfied*, 3 = *mixed feelings*, 4 = *satisfied*, 5 = *very satisfied*. The questions were identical to those mothers completed initially.

Mothers were asked to recall the helpfulness of *social support* received from 7 different sources when their children were 5 months old: (1) child's father, (2) mother's family and relatives, (3) father's family and relatives, (4) child's pediatrician, (5) friends or neighbors, (6) an organized course, class, or group, and (7) books, magazines, or newspapers. Each item was rated on a 5-point scale from 0 (*not used* or *of little or no help*) to 4 (*very helpful*).

Mothers were asked to reflect on four dimensions of their parenting role: satisfaction, competence, investment, and role balance. The Self-Perceptions of the Parental Role (SPPR; MacPhee, Benson, & Bullock, 1986) was re-administered verbatim. The statements were written in the present tense, and mothers were instructed to think back to how they saw themselves as parents when their children were 20 months old. Each item had a pair of statements that described contrasting endpoints of the dimension in question, thereby minimizing socially desirable responses. For example, one of the items stated: "Some parents do a lot of reading about how to be a good parent. BUT Other parents don't spend much time reading about parenting." The mother chose the statement that described her best and then checked *Sort of true for me* or *Really true for me*. There were four response items, weighted 1, 2, 4, and 5 to account for the absence of a response indicating that the item was equally like and unlike the mother. Scores for each scale were the mean of responses to the items comprising the scale.

Mothers were asked to recall their beliefs about the importance of their influence on their 48-month-old children's social skills, knowledge, independence/initiative, creativity, playfulness, planfulness, confidence, self-control, morals and values, providing companionship/attention, and providing variety of experience. The response scale ranged from (1) *not important* to (6) *very important*. This scale was administered unchanged.

Domain 2 Measures: Mother-Reported Child and Mother Behaviors

Mother-reported child behaviors.

Maternal memory of child behavior was assessed using multiple measures. Mothers were asked about the child's *birth weight* in pounds and ounces and *birth length* in inches. Responses were then converted to grams and centimeters, respectively. Mothers were also asked about the baby's *maturity at birth* in terms of whether the child was born preterm (more than 3 weeks before due date), term (within 3 weeks of due date), or postterm (more than 3 weeks after due date). These questions were originally gathered at the longitudinal study intake interview just prior to the 5-month visit.

Several questions asked about mothers' perceptions of the *baby's adjustment in the first month of life*. One was a global rating of the baby's ability to ''settle in'' and adjust to feeding, sleeping, people, and routines, rated on a scale from (1) *very easy adjustment* to (5) *very difficult adjustment*. Another 10 questions were included to assess common problem

behaviors experienced by infants. Mothers checked whether (1) or not (0) the baby had experienced each of these behaviors in the 1st month at home: excessive vomiting, prolonged or frequent diarrhea or constipation, pronounced lack of interest in being fed or active refusal to eat, excessive demand to be fed, frequent waking and crying at night, excessive sleeping during the day, frequent and intense crying generally, lack of interest in things going on around him or her when awake and alert, noticeable stiffening, turning away, or crying when picked up or handled, and pronounced clinging when picked up or intense crying when put down. Responses to the 10 items were reversed and summed to create a scale of mothers' perceptions of the Absence of Difficult Behaviors in the First Month. These questions were administered verbatim from the original versions asked at the 5-month visit.

The Vineland Adaptive Behavior Scales Interview Edition, Survey Form (VABS; Sparrow et al., 1984) were used to obtain mothers' estimates of their 20- and 48-month-old children's adaptive behaviors in four domains: (1) receptive, expressive, and written communication skills, (2) daily living skills, (3) social skills, and (4) large and small muscle motor skills. The interview assesses the child's performance. The VABS Survey Form was normed on 3000 individuals aged newborn to 18 years and 11 months, and its validity is well supported (Sparrow et al., 1984). For mothers' memories of children's performance at 20 and 48 months, we selected 6 questions from each domain to represent items that were *easy*, *mid-range*, and *difficult* for the age group based on children's ratings at baseline. Mothers checked (2) *Yes* if their child was definitely doing the behavior, (1) *Learning* if the child was developing the behavior, or (0) *No* if the child was not yet doing the behavior at 20 and at 48 months. The 6 items for each domain were averaged to compute a total domain score at each age.

Mothers rated their 48-month-old child's *degree of enjoyment* of each of 13 activities, such as play, outings, conversation, watching TV, formal education, chores, and mealtimes, on a scale from (1) *not at all enjoyable* to (6) *very enjoyable*. Items were averaged to create a single scale of the child's enjoyment.

Mothers indicated, on a scale from (1) *not well* to (6) *very well*, how well they thought their children would have done on 11 activities at 48 months. The activities were a game with rules, a game with numbers, drawing a person, drawing in general, telling a story, verbal expressiveness, self-understanding, interpersonal understanding, social skills, block building, creativity, and physical coordination. Ratings for all 11 items were averaged to create a scale of mothers' Expectations of Child Performance.

The Preschool Behavior Questionnaire (PBQ; Behar & Stringfield, 1974) was completed by mothers when their children were 48 months to detect problem behaviors. The questionnaire was re-administered in the same form, and mothers were asked to recall their children's behaviors at 48 months of age. Each of 30 PBQ questions was rated on a 3-point scale for the degree to which a target behavior applied to the child: (0) *Never*, (1) S*ometimes*, and (2) *Often*. All items were reverse-coded, and scales were computed to reflect the absence of problem behaviors in three areas: non-hostile/aggressive, non-anxious/fearful, and non-hyperactive/distractible.

Mother-reported mother behaviors.

Mothers' memories of their own behaviors consisted of questions about the pregnancy and birth, hours of employment, and parenting practices. All questions were re-administered unchanged. Mothers were asked to indicate (1) whether or (0) not they had *planned the pregnancy*, experienced any *problems or complications in the pregnancy*, and experienced any *problems or complications in the labor and delivery*. The frequency of prenatal care (6-point scale ranging from 0 = *never* to 5 = *once/month or more often*) was also assessed.

Mothers listed their employment histories by the age of their child, including hours of employment. From this list, we extracted mothers' hours of employment when her child was 5, 20, and 48 months of age. The original data were collected in the same manner.

A 16-item self-report measure of mothers' behaviors toward their 20-month-olds, the Parental Style Questionnaire (PSQ; Bornstein et al., 1996), asked mothers to rate on a scale from (1) *hardly at all* to (5) *all the time* how much they engaged in specific parenting behaviors (e.g., "I provide my child with positive affectionate displays of warmth and attention."). Items formed three domains: *social exchange* (4 items), *didactic interaction* (9 items), and *limit setting* (2 items). The PSQ subscales have demonstrated construct validity and internal consistency (Bornstein et al., 1996). This questionnaire was re-administered unchanged.

Domain 3 Measures: Observed Child and Mother Behaviors

Observed child behaviors.

Assessment of maternal memory of observed child behaviors was based on data collected during videorecorded home visits, where either everyday naturalistic behavior was observed for 1 hr (at 5-months) or mother-child free play was observed for 10 min (at 20 months).

To assess maternal memory of 5-month child behaviors, we asked questions about the frequencies of specific behaviors. The domains included vocalizing, sitting, and moving. Mothers were asked how often the baby *babbled or cooed* and how often the baby *fussed or cried* at 5 months. The scale ranged from (0) *never or almost never* to (4) *very often*. Mothers also stated whether their child had achieved several sitting and movement milestones at 5 months. *Sitting* levels ranged from sitting upright only with full supports to being able to rotate from an unsupported sitting position to hand-knee or tummy without assistance. *Movement* levels ranged from lifting up legs while lying on the back to well-coordinated and controlled hand-knee crawling. The highest levels of sitting and moving that the mother indicated were compared to the highest levels of sitting and moving that were observed and recorded in the hour of interaction at 5 months.

Children's language comprehension and production were measured at 20 months with the Vocabulary subscale of the Expressive Language Scale and the Verbal Comprehension Scale 'A' of the Reynell Developmental Language Scales--Second Revision (RDLS; Reynell & Gruber, 1990). The RDLS is a standardized instrument for assessing language expression and comprehension among children aged 1 to 6. The RDLS was originally administered by

an experimenter in the presence of the mother, but the mother was not an active participant. We extracted and reworded questions from 4 sections of the RDLS to assess mothers' recall of their children's verbal comprehension and verbal production. Mothers were asked to (1) identify objects the child could point to when named (e.g., ball, spoon), (2) indicate whether the child could follow 2-step directions (e.g., put the doll on the chair), (3) indicate the highest developmental level of spoken language (ranging from babbling, to word combinations, to appropriate use of pronouns), and (4) indicate which words the child could speak when presented with an object (e.g., ball, spoon).

Observed mother behaviors.

Mothers' memories of their own observed behaviors were also based on videorecorded data at 5 and 20 months. We asked mothers to recall how often they talked to their 5-month-old babies, engaged in social play (e.g., peek-a-boo), encouraged attention to themselves, and encouraged attention to properties, objects, or events in the environment. These 4 questions were rated on a scale from (0) *never or almost never* to (4) *very often*. Mothers also characterized their global style of directing the baby's attention on a continuum from (0) *more likely to engage the baby in social exchanges* to (7) *more likely to direct the baby's attention to objects*.

To map onto play behavior when children were 20 months, mothers were asked to recollect their frequency of play demonstrations (e.g., building a tower), play solicitations (e.g., suggesting the child throw the ball), and social play (e.g., playing This Little Piggy) when their children were 20 months old. These 3 questions were also rated on a scale from (0) *never or almost never* to (4) *very often*.

Confidence and Explanatory Measures

Maternal confidence in her memory.

Mothers were asked to indicate how confident they were in the consistency of their memories about each of the following: the child at 5 months, 20 months, and 48 months, herself at 5 months, 20 months, and 48 months, employment history, childcare, and social support. Mothers indicated their confidence by slashing a 7-cm semantic differential line that ranged from *non-confident* to *confident*, and scores were coded into half points on a scale from 0 to 7.

Explanatory variables.

We used several variables about the mother, child, and situation to explore relations with consistency of mothers' memories. Maternal age, education (on the 7-point Hollingshead, 1975, scale), hours of employment, child gender, number of siblings, and the recall assessment interval were all gathered from the demographic questionnaire. In addition, mothers had completed a personality inventory and social desirability scale when their children were 14 years, and a measure of verbal intelligence when children were 20 months.

The Jackson Personality Inventory-Revised (JPI-R; Jackson, 1994) was developed to provide measures of personality likely to have a significant impact on a person's functioning. The

revised version contains 300 *True-False* statements, representing 15 personality traits which may be grouped into five meaningful clusters that correspond with the Big Five (Digman, 1990). Following the model outlined by Jackson (1994), the personality dimension of Neuroticism was estimated as the principal component of empathy, anxiety, and cooperativeness; Openness was estimated as the principal component of complexity, interest, innovation, and tolerance; Extraversion was estimated as the principal component of sociability and confidence; Conscientiousness was estimated as the principal component of organization and responsibility; and social astuteness was used as the single indicator of Agreeableness.

The short form of the Social Desirability Scale (SDS-SF; Reynolds, 1982) uses 13 of the original 33 items in Crowne and Marlowe's (1960) Social Desirability Scale (SDS) to assess adults' tendencies to respond to questions in a socially desirable fashion. Statements like, "I'm always willing to admit when I make a mistake" are rated as *True* or *False*. Reliability of the SDS-SF is .76, and the correlation with the full-length SDS is .93 (Reynolds, 1982).

The Peabody Picture Vocabulary Test-Revised (PPVT-R Form L; Dunn & Dunn, 1981) presents up to 175 vocabulary words verbally by a trained administrator, and for each word presented the mother chose one of four pictures to indicate the meaning of the word. Standard scores with a possible range of 40 to 160 ($M = 100$, $SD = 15$) were obtained based on the mothers' age. The median split-half reliability coefficient for 828 adults ranging from 19 to 40 years was .82.

Analysis Plan

To assess the consistency between the original responses and performance of children and mothers with mothers' memories of those responses and performance, we aggregated the original and recalled data in the same ways. For example, we used the same questions from each domain of the VABS to compare original and recalled data. In most cases, this strategy resulted in scales or items that were directly comparable for determining mean level and relative standing. However, in Domain 3 we were comparing observed behaviors to mothers' memories of the frequencies of those behaviors. For those items, we were not able to equilibrate the scales so that the means could be compared. Therefore, we used only agreement in relative standing.

In presenting the results of each domain, we organize the tables by age group. For each item or scale, we computed paired-samples *t*-tests to assess consistency in mean levels and Pearson's correlations to assess consistency in relative standing. (We did not use the intraclass correlation, which standardizes variables across measurements, because some items were measured slightly differently over time.) If the item being compared was categorical, we calculated chi-squares with follow-up configural frequency analyses (CFA; von Eye, 2002) and percent agreement instead of a *t*-test and correlation, respectively. A first-order CFA using Lehmacher's test with Küchenhoff's continuity correction indicates whether each cell in a cross-tabulation contains more or fewer cases than would be expected based on the marginal frequencies. If consistency is good, we expected a significant and positive z-score, indicating that proportionally more cases fell on the diagonals than the off-

diagonals of the cross-tabulation. For each section in Tables 1–3, we also computed summary statistics by age group. Summary statistics were the average *t*-test or chi-square and the average correlation (using Fisher's *Z*-transformation; Glass & Hopkins, 1996) or percent agreement. No significance tests were available for these summary statistics, but for purposes of discussion, we considered an average absolute *t*-value over 2.00 and an average absolute correlation of .30 to be significant (based on a sample size of 46). For all tests, we screened for influential outliers and, if found, report the results with and without them (Fox, 1997).

For mean-level consistency, we considered a significant *t*-value to indicate poor consistency and a nonsignificant *t*-value to indicate consistency. In discussing consistency in relative standing, we followed Cohen's (1988) conventions for effect size, considering a correlation of .10 to indicate low consistency, .30 moderate consistency, and .50 strong consistency. We could not find any published conventions for percent agreement, so to have equal intervals we chose 25% for low consistency, 50% for medium consistency, and 75% for strong consistency.

A post-hoc power analysis was computed prior to data analysis to determine whether the sample size of 46 provided sufficient power to detect a medium-sized effect in paired *t*-test and correlation analyses. With a medium effect size (d_z = .50 for the paired *t*-test, and r = .30 for the correlation; Faul, Erdfelder, Buchner, & Lang, 2009), α = .05, and N = 46, the power estimate was .91 for the paired *t*-test and .55 for the correlation. Testing the same parameters with a large effect size (r = .50) produced a power estimate of .97 for a correlation analysis. These power estimates indicate adequate power to detect a medium or large effect in paired *t*-tests and a large effect in correlation tests.

To assess relations between consistency and mothers' confidence in their memories, we correlated maternal confidence ratings for each domain with the absolute discrepancy in scores between the original data and the recalled data in that domain. We used absolute discrepancy scores because we were interested in consistency of maternal memories independent of the direction of difference (i.e., recalling more or less of a certain behavior). Likewise, to determine whether maternal, child, and situational characteristics were associated with consistency in original and retrospective reports, we correlated these characteristics with absolute discrepancy scores.

Results

Domain 1: Maternal Cognitions

Table 1 displays the results for Domain 1, on consistencies of mothers' memories of their own cognitions when their children were 5, 20, and 48 months.

At 5 months, mothers recalled cognitions as more positively colored than they originally reported (higher scores on 5 of 9 items), but they were moderately consistent in their relative standing on cognitions as when their children were 5 months of age (significant consistency on 6 of 9 items). At 20 months, there was no difference between mothers' average recall and their cognitions when their children were 20 months of age (except that mothers

remembered being less satisfied than they originally reported), and mothers' relative standing was moderately consistent with that of the 20-month reports (significant consistency on 3 of 4 items). At 48 months, mothers remembered being less satisfied with their childcare arrangements and having lower perceived influence over their children's development than they originally reported at 48 months. As at 5 months, mothers reported similar relative standing on their satisfaction with the balance between parenthood and employment, but not their childcare arrangements or influence on the child at 48 months.

Domain 2: Mother-Reported Child and Mother Behaviors

Table 2 summarizes the results for Domain 2, on consistencies in mothers' memories of their reports of their children's and their own behaviors when their children were 5, 20, and 48 months.

Mother-reported child behaviors.

At 5 months, there was no difference in average recall of mothers' reports of their children's behaviors when their children were 5 months of age and nearly 16 years later on any item. Mothers were also strongly consistent in relative standing in the recalled and original data (significant consistency on all 5 items). At 20 months, mothers tended to recall their children's adaptive behaviors more positively than they originally reported (higher scores on 2 of 4 items) and did not maintain similar relative standing (significant consistency only on VABS communication). At 48 months, mothers again tended to recall their children's behavior as more positive than they originally reported (higher scores on 7 of 10 items; lower score on 1 of 10 items), and their relative standing on recalled reports of their children's behaviors was consistent with that reported at 48 months (significant consistency on 8 of 10 items).

Mother-reported mother behaviors.

At 5 months, over 65% of mothers correctly recalled their original answers about whether they had planned the pregnancy, had pregnancy or birth complications, and received regular prenatal care. However, configural frequency analyses indicated that mothers were not consistent at remembering whether they had planned the pregnancy, $z = 1.46$, *ns*, but were consistent at recalling pregnancy complications, $z = 13.53$, $p \leq .001$, and birthing complications, $z = 2.48$, $p \leq .01$. Mothers reported at least moderate consistency in relative standing on all items. At 20 months, there was little difference in mothers' average recall of their own behaviors (except that mothers remembered less social exchange than they originally reported), and mothers reported consistent relative standing to their original reports (significant consistency on all 4 items). At 48 months, mothers consistently reported their reported hours of employment at 48 months (e.g., similar mean and high correlation).

Domain 3: Observed Child and Mother Behaviors

Table 3 shows the results for Domain 3, on consistency in mothers' memories of their children's and their own observed behaviors when their children were 5, 20, and 48 months.

Observed child behaviors.

Mothers estimated their children's behaviors as more positive than was observed at 5 months (higher scores on both behaviors), and they did not report similar relative standing to what was observed at 5 months (significant consistency only on the highest level of movement). At 20 months, there was no difference in mothers' average estimations and the original observations (except that mothers overestimated how many 2-step directions their children could follow), but mothers' reports did not have similar relative standing to any observed behavior.

Observed mother behaviors.

At 5 months and 20 months, the relative standing of mothers' judgments was not consistent with ratings of any mother behaviors in the naturalistic observation with their child.

Summary Views

Summary by domain.

Across all maternal cognitions we measured in *Domain 1*, mothers reported similar average levels (i.e., nonsignificant *t*) on 8 of the 16 cognitions (50%), recalled higher levels (i.e., significant, positive *t*) on 5 cognitions (31%), and recalled lower levels (i.e., significant, negative *t*) on 3 cognitions (19%). Mothers also reported similar relative standing (i.e., significant, positive *r*) on 10 of the 16 cognitions (63%) we measured.

Across all the behaviors we measured in *Domain 2*, mothers reported similar average levels on 13 of the 27 behaviors (48%) we could test, recalled higher levels on 12 behaviors (44%), and recalled lower levels on 2 behaviors (7%). Mothers also reported similar relative standing to their original reports of their own and their children's behaviors on 21 of the 29 behaviors (72%).

Across all the behaviors we observed in *Domain 3*, mothers estimated similar average levels on 3 of the 6 behaviors (50%) we could test, recalled higher levels on 3 behaviors (50%), and recalled lower levels on 0 behaviors (0%). Mothers also had similar relative standing on only 1 of the 16 behaviors (6%) we measured.

Summary by age of child.

Across all the cognitions and behaviors we measured at *5 months*, mothers recalled similar average levels on 9 of the 19 behaviors (47%) we could test, recalled higher levels on 10 behaviors (53%), and recalled lower levels on 0 behaviors (0%). Mothers also had similar relative standing on 14 of the 28 items (50%) we measured.

Across all the cognitions and behaviors we measured at *20 months*, mothers estimated similar average levels on 11 of the 16 behaviors (69%) we could test, recalled higher levels on 3 behaviors (19%), and recalled lower levels on 2 behaviors (12%). Mothers also had similar relative standing on 8 of the 19 items (42%) we measured.

Across all the cognitions and behaviors we measured at *48 months*, mothers recalled similar average levels on 4 of the 14 behaviors (29%), recalled higher levels on 7 behaviors (50%)

and recalled lower levels on 3 behaviors (21%). Mothers also had similar relative standing on 10 of the 14 items (71%) we measured.

Summary by participant.

For mothers' ratings of the *child* across all domains and ages, mothers recalled similar average levels on 11 of the 24 behaviors (46%), recalled higher levels on 12 behaviors (50%), and recalled lower levels on 1 behavior (4%). Mothers also had similar relative standing on 14 of the 27 items (52%) we measured.

For mothers' ratings of *mothers'* own behaviors and cognitions across all domains and ages, mothers recalled similar average levels on 13 of the 25 behaviors (52%), recalled higher levels on 8 behaviors (32%), and recalled lower levels on 4 behaviors (16%). Mothers also had similar relative standing on 20 of the 34 items (59%) we measured.

Summary by metric.

Across all domains, ages, and participants, mothers recalled similar average levels on 24 of 49 behaviors (49%), recalled higher levels on 20 behaviors (41%), and recalled lower levels on 5 behaviors (10%). Mothers also had similar relative standing on 32 of the 61 items (52%) we measured.

Maternal Confidence in Her Memory

Mothers were asked to indicate their confidence in their memories in several areas. Table 4 displays the descriptive statistics for mothers' confidence ratings. Of 45 correlations between maternal confidence ratings and absolute discrepancy in scores between the original data and the recalled data, only 5 were significant (11%), and no clear pattern emerged. Maternal confidence seems to have little association with the consistency of maternal memories.

Understanding Maternal Memories

We investigated whether maternal, child, and situational characteristics were related to the absolute discrepancy between mothers' original reports and their retrospective reports.

Maternal characteristics.

We explored the possibility that maternal personality characteristics, social desirability bias, verbal intelligence, age, education, and hours of employment were associated with consistency of maternal memories. With a few scattered exceptions, these characteristics were not systematically associated with consistency. Only 25 of 450 tests (6%) were significant, and the pattern was unsystematic.

Child characteristics.

We explored whether mothers of girls and boys had similar absolute discrepancy scores, and with 2 exceptions of 45 tests (4%) they did. We also considered the possibility that having more children in the family affects how precise mothers' memories were. With one

exception in 45 tests (2%), the number of children in the family was unrelated to the absolute discrepancy of reports.

Situational characteristics.

To determine whether variation in time between original data collection and the retrospective survey related to maternal reports (15, 14, or 12 years), we correlated the number of years between assessments with the absolute differences in scores at the original and retrospective data collection points. With one exception of 45 tests (2%), the time between assessments was unrelated to maternal memories.

Discussion

Parent retrospective reporting has a long and checkered history in developmental science. The present study compared retrospective to original reports collected from mothers in a within-mother multi-wave longitudinal design. The study offers several new points of clarification to the debate about the reliability and validity of maternal retrospective report. First, this was not a sterile laboratory study of long-term recollection of arbitrary stimuli. We collected and compared maternal reports about personally and ecologically meaningful information across three different domains: maternal cognitions, mother-reported child and mother behaviors, and observed and coded child and mother behaviors that included data on cognitions, practices, and behaviors and on children and mothers themselves. Autobiographical memories (such as we studied) play significant roles in our lives – they contribute to the sense of self, regulate emotions, create and maintain social ties, and help us teach and inform others – and do so in ways that laboratory tasks cannot match (Bluck, 2003). Second, we used a repeated-measures design spanning from infancy to adolescence and compared three different long-term time points from 5, 20, and 48 months to 16 years. Third, we evaluated consistency in maternal reports from both mean-level and rank-order vantages. Fourth, we brought to bear several variables about the mother, child, and situation to potentially explain consistency of mothers' memories. Fifth, we collected confidence ratings from mothers about their retrospective reports. Sixth and last, we re-administered the questionnaires completed by mothers at the three original time points unchanged (or with only minor changes) from their original form. The assessment of retrospective recall necessarily relies on the quality of the measures of experiences at both time points and on the comparability of the measurement at the original time and as retrospectively reported later. Our measures ensured comparability in these respects. Many studies fall down on this criterion because the measures are vague or global or noncomparable. Notably, in their classic study on which so much reputation of maternal retrospective report has turned Yarrow et al. (1970) relied on nursery school records deriving from unstandardized parental reports.

Overall, we found that mothers recalled similar or higher levels of most cognitions and behaviors examined. Mothers also had similar relative standing on just over half of items measured. Retrospective parental report represents a reasonably valid method for studying child development and parenting. Several other studies suggest that retrospective reports can provide valuable information, assuming possible biases and limitations of the method are

addressed (Bernstein et al., 1994; Brewin et al., 1993; Hardt & Rutter, 2004; Henry, Moffitt, Caspi, Langley, & Silva, 1994; Pickett, Kasza, Biesecker, Wright, & Wakschlag, 2009). Even Yarrow et al. (1970, p. 69) found significant overlap between baseline and recall in their original study, conceding that retrospective reports "may reflect to some extent what relations exist." The takeaway is that retrospective studies continue to occupy a worthwhile place in developmental research, but further research is needed to examine possible biases in reporting (Hardt & Rutter, 2004). Generally, memory research focuses on the overall accuracy/inaccuracy of memories instead of the factors that foster or impede retrospection.

Our data indicate that some retrospective parental reports should continue to be interpreted with caution. First, even large relative standing (*r*) consistency of .50 means that retrospective and original scores share only a quarter of their common variance. In our data, only reports of birth weight, hours of employment, and PSQ social exchange shared more than 50% of their variance. Several possible explanations for why retrospective recall may sometimes fail have been advanced, including the dependence of recall on attention and encoding, the vagaries of memory that lead to forgetting, retroactive interference based on what has happened subsequent to initial encoding, the tendency to impose meaning on memories, and mood states at the times of encoding or retrospection (Bower, 1981; McFarland & Buehler, 1998; Schacter, 2007; Schacter, Koutstaal, & Norman, 1996). Modest levels of consistency may instill consternation, but should be interpreted in context. In the behavioral sciences, modest levels of agreement are *de rigeur* even between different reporters (Bornstein et al., 1991). It is possible in such circumstances to interpret one reporter's information as invalid, but it is also possible to interpret this finding as reflecting that different reporters have access to different and valid but complementary information. On this account, it is therefore possible that prospective and retrospective reports share some things in common but also contain unique information. That is, modest agreement may indicate that retrospective and prospective measures are better regarded as complementary, rather than as more or less valid. Indeed, it could be argued that in certain conditions retrospective reports hold more validity than in-the-moment reports because they place the parent-child dyad in a broader developmental context. For example, a parent might not recall child behaviors that once seemed dysfunctional (e.g., biting or hitting other children) because they were transient, but another parent may well remember the same behaviors because they were characteristic of a pattern of aggression that persisted. This phenomenon is illustrated by the "response shift bias" (Howard, 1980), where the person's evaluation standard changes as a result of experience. Because of the response shift bias, intervention studies frequently employ a retrospective pretest (Hill & Betz, 2005; Pratt, McGuigan, & Katzev, 2000). In essence, parents might not have the benefit of experience with their child when they are assessed at a particular stage, but when completing retrospective reports they possess the wisdom that comes with parenting experiences garnered in the intervening time.

Second, the consistency of retrospective reports in this study varied by domain, child age, person, and metric of measurement. Not all domains are remembered equally well. This result should be unsurprising, as it is actually typical of recall. For example, higher-priority items are recalled better than lower-priority items, even when the two are presented at the same time (Sakaki et al., 2014): Memory for an action per se is enhanced by emotion, but

emotion does not enhance memory for which person performed which action, and memory for negative events is often better than memory for positive events (Kensinger, 2009).

Third, parenting is an inherently emotional undertaking (Dix, 1991; Leerkes & Augustine, in press). The subject of memories here – mothers' own infants and children and how those mothers parented them early in their lives now that they are adolescents – unquestionably taps emotions. One important aspect of human memory concerns binding contextual information (e.g., time, place, or associative cues) that constitutes many disparate features of a unified event (Davachi, 2006; Ranganath, 2010). Extant evidence identifies both enhancing and impairing effects of emotion on memory binding, and these opposing effects have been interpreted in the context of different (but not mutually exclusive) views (Bisby, Horner, Hørlyck, & Burgess, 2016; Chiu, Dolcos, Gonsalves, & Cohen, 2013; Christianson, 1992; Kensinger, 2009; Mather & Sutherland, 2011). For example, emotional arousal can narrow attention to central cues at the expense of peripheral and irrelevant cues (Easterbrook, 1959). Such prioritization of resources toward central aspects of information can enhance memory for those aspects compared to peripheral or contextual ones that are attended less (Buchanan & Adolphs, 2002; Burke et al., 1992; Christianson, 1992; Kensinger, 2009). Mather (2007) suggested that, when objects are emotionally arousing, attention is focused more intensely on those objects, and thus emotional arousal can increase the binding of the objects in memory.

Mothers retrospectively reported more positive cognitions than they reported at 5 months, more positive child behaviors than at 20 and 48 months, and more positive maternal behaviors than at 5 months. These findings support extant literature indicating that mothers' reports about the past tilt towards positive recollections (Kagan, 1998). Yarrow et al. (1970), too, reported that mothers recall their own roles as mothers positively. More generally, this pattern of results comports with the age-related "positivity effect" (i.e., the tendency to enhance positive emotions and reduce negative emotions; Mather, 2016). This so-called "Pollyanna principle" (Matlin & Stang, 1978), and the related phenomenon of "fading affect bias" (in which the strength of emotions associated with negative autobiographical memories fades more rapidly than it does for positive ones; Walker & Skowronski, 2009) have been argued to be psychologically protective, to keep one's emotional life pleasant (Kappes & Crockett, 2016; Kouchaki & Gino, 2016; Richard, Skowronski, & Thompson, 2003; Ritchie et al., 2006) and to guide current and future behavior. A number of other salutary consequences follow from such optimizing memory reconstructions. Memory is used to remember the past, but it is also applied to interpret the present and to anticipate the future. So, how we view what we remember is important, and future parenting can take advantage of the positive-event recall bias (Ros & Latorre, 2010).

Mothers overestimated their children's developmental levels of sitting and moving at 5 months, how many 2-step directions their children could follow at 20 months, and their children's skills at 48 months. This pattern contradicts older research stating that parents underestimate their children's abilities retrospectively (D'Alessio, 1990; DeGrada & Ponzo, 1971); it seems rather that parents tend to paint their own and their children's past in a positive light. In this connection, however, we did not find mothers' social desirability bias was related to consistency of their reports. Future research might profitably take up the

question of why retrospective parental reporting is rosy. Hints might come from Robbins (1963), who found that inaccuracies in parents' recall of their childrearing practices went in the direction of expert recommendations, and from DeGrada and Ponzo (1971), who thought that parents tend to revert to stereotypes when reporting retrospectively on children's development.

Previous research has implicated several non-cognitive factors in the consistency of retrospective reports. For example, parental report consistency may be influenced by maternal and child characteristics, such as age, personality, gender, education, family size, and SES (Bornstein, 2014; Bugental & Corpuz, in press; De Los Reyes & Kazdin, 2005; Sameroff et al., 1982; Vaughn et al., 1987). Our findings did not point to meaningful relations with any of several child or maternal or situational characteristics on maternal report consistency, with only unsystematic chance-level relations emerging. This result agrees with previous research which assessed parental recall of childhood illness and found that caregiver educational attainment and occupation were not associated with accuracy of recall (Pless & Pless, 1995). Associations between personality factors and the propensity to recall are also extremely modest (Reuben et al., 2016). However, we did not assess all possible individual-difference factors. For example, people with high and low working memory capacity show variability on various tests of episodic long-term memory including recognition (e.g., Unsworth, 2010; Unsworth, Brewer, & Spillers, 2011).

One factor not studied here, but potentially influential in maternal retrospective reports about early child development and parenting, is stress. Memory is influenced by stress and its associated neuroendocrine responses. Stress can enhance memory encoding and consolidation, but stress can also interfere with memory and so induce forgetting (Lupien et al., 1997). Recall performance following a stressful event is worse than after a non-stressful event. Stress biases attention and prefrontal cortical function, which could further affect controlled retrieval. Similarly, common in mothers after birth, depression has been associated with a tendency to recall more negative experiences (Whalley, Rugg, & Brewin, 2012).

When people express high confidence in their memories, they are likely to be more accurate: High confidence heralds high accuracy in recognition. Mothers in this study tended to be confident in their recollections (e.g., means in the top half of the scale; Table 4), but we found little evidence for any relation between maternal confidence and the consistency of maternal retrospective memories with original events. Memory researchers have learned (the hard way) that relations between confidence and accuracy are never perfect. Within certain limits, however, the degree of certainty in one's replies should bear a relation to the fidelity of those replies (Dallenbach, 1913). Wixted and Wells (2017) argued that one should expect a strong relation between confidence and accuracy when conditions are pristine. We asked about confidence for sets of items rather than individual items, which may explain why we failed to find a close match between confidence and consistency of reports.

We also found the time between assessments essentially a non-factor in recall. Some memory researchers have observed that forgetting naturally increases over longer retention intervals, whereas others have observed that longer retention intervals facilitate memory

consolidation (Dolcos et al., 2005, 2012, 2017; LaBar & Cabeza, 2006; Ritchey et al., 2008; Weymar et al., 2009, 2011; Weymar & Hamm, 2013; Yonelinas & Ritchey, 2015). In accord with our results, Yarrow et al. (1970) reported that inter-assessment time did not consistently influence recall of life domains.

As the explanatory variables in the present study did not impact the consistency of retrospective reports either, future research should examine additional possible factors. The ways in which participants are asked about experiences likely contribute to the validity of their reports. For example, report format and question wording have been shown to affect retrospective reports (Barsky, 2002; Schwarz, 1999; Schwarz & Oyserman, 2001). Retrospective report validity can also be improved by asking reporters for specific rather than general information (Yarrow et al., 1970). Because the consistency of retrospective reports in our study varied by domain, it is also likely that the content of questionnaires systematically influences mothers' consistency. Mothers' parenting knowledge, attitudes, and moods at the times of original experience and recall may influence consistency (Barsky, 2002; Bower, 1981; Lewinsohn & Rosenbaum, 1987; Seifer et al., 1994). Mothers have also been reported to distort their reports to match their current perceptions of their children (Chess et al., 1966; Robbins, 1963; Yarrow et al., 1970). It is possible that our study findings were influenced by the format of our questionnaires. Memory depends on encoding, storage, and retrieval of material. Our study was not designed to identify where in the memory stream recollections for parental information consolidate or falter. Future research should be designed to do so.

Future research will further benefit from larger and more diverse samples. The mothers in the present study were all European American and predominantly educated. Ours was a socioeconomically heterogeneous (not homogenous middle-class) community sample in terms of maternal education and family SES, but an racially homogenous sample. European Americans are a large and possibly diverse racial group in the United States in themselves (Humes, Jones, & Ramirez, 2011; Tilton-Weaver & Kakihara, 2008), originating from disparate parts of the world. Child behavior and parenting processes alike are known to vary with race and ethnicity (e.g., Graham, 1992; Parke, 2000; Tomlinson, Bornstein, Marlow, & Swartz, 2014). Even our relatively homogenous ethnic sample therefore constitutes only a first step in understanding the matrix of associations surrounding recollections of child development and maternal parenting cognitions and practices that logically antecedes embarking on more complex studies and analyses with diverse samples. However, this study intentionally avoids major racial confounds that have plagued the parenting literature and would cloud our findings (Bornstein, Jager, & Putnick, 2013; Jager, Putnick, & Bornstein, 2017). Nonetheless, future research should be conducted comparing ethnic and racial groups, SES, and cultures to identify the specificity and expand the generalizability of the results.

The present study included maternal reports only. Extant evidence points to sex differences in emotion processing (Andreano, Dickerson, & Barrett, 2014; Stevens & Hamann, 2012) and in emotional memory (Andreano & Cahill, 2009; Cahill, 2003; Hamann, 2005). For example, compared to men's, women's emotional memories are enhanced (e.g., Canli et al., 2002; Davis, 1999; Seidlitz & Deiner, 1998), and women show enhanced overall brain

response during encoding of emotional memories (Canli et al., 2002). We focused on mothers due to their central role in childrearing. Mothers recall health events more accurately than fathers (Pless & Pless, 1995). We expect that fathers would perform worse than mothers on these measures. Jager, Bornstein, Putnick, and Hendricks (2012) showed that each individual in the family owns a unique perspective on family functioning. Future research should enlist reporters other than mothers, minimally comparing maternal and paternal reports. That is, gender differences should figure in future research designs and analyses (Sakaki et al., 2012).

The "stimuli" we used were high in ecological validity, but in using them we relinquished some experimental control: The events we asked mothers to recall may follow similar scripts but differ in terms of many details from one mother to another. We had no control over whether mothers previously accessed the information we asked about or how often they did. Experimentally, people recall a higher proportion of items that they had previously retrieved relative to items that were untested (the "retrieval practice effect"; Sutterer & Edward, 2016).

Finally, future research could compete original and retrospective reports to ascertain their relative predictive validity. For example, if contemporary reports of problem behaviors predict later child adjustment, that analysis could be repeated with retrospective reports of problem behaviors to determine if retrospective reports are equally, more, or less predictive. It is possible that knowing how the child developed informs retrospective reports and improves predictive relations to later functioning over concurrent reports. Furthermore, concurrent child characteristics could be controlled to ascertain if parents use the child's current functioning to "remember" past functioning.

Implications and Conclusions

Overall, our findings suggest that retrospective parental reports have a place in developmental science and could continue to be used albeit with some caution. Multiple domains and multiple indicators should be used for each domain in place of single domains and single-item indicators. Retrospective reports may also be more accurate as questionnaires or checklists, as these formats rely less on recall memory than recognition (Bornstein & Toole, 2010). Retrospective report validity can also be improved by asking reporters for specific rather than general information (Yarrow et al., 1970). Retrospective parental reports should also be combined with objective observations and testing whenever feasible. To capture the nuances of child development and parenting, researchers should collect data from multiple reporters and use multiple methodologies (Bornstein, 2014). Study goals should be weighed in terms of the advantages and disadvantages of retrospective reports. For example, if participants are classified into groups based on reports of preschool problem behaviors, retrospective reports are likely to produce different groups than original reports because of the positive bias in retrospective reports. However, if predictive relations will be explored between preschool problem behaviors and later functioning, retrospective reports may be adequate because consistency in relative standing between original and retrospective reports tends to be high.

To enhance recall during interviews, a mixture of open-ended and more systematic questioning has proven effective (Rutter et al., 1998). Respondent-based interviews use highly structured questionnaires with precise wording and yes/no responses. Investigator-based interviews gather comprehensive descriptions of the behaviors and details surrounding an event. Combining these two methods may improve retrospective reports. The cognitive interview, which amalgamates the two methods using principles from cognitive psychology, enhances recall in eye-witness reports (Saywitz, Geiselman, & Bornstein, 1992). Some methods, such as the Life History Calendar and the Event History Calendar, also yield reliable retrospective data. The Life History Calendar (LHC; Caspi et al., 1996) uses visual aids by creating large grids to record streams of events in the reporter's life. The Event History Calendar (EHC; Belli, Shay, & Stafford, 2001) collects data in the form of time lines and is thought to enhance recall for events occurring years in the past. Both methods use a process called "personalized timing" to anchor the record of events through a series of temporal landmarks.

Parental reports are fundamental in child development and family research, and they are likely to remain so. The validity of such reports has important implications for the field at large. Because the consistency of reports may vary, further research is needed to determine the most appropriate applications and interpretations of the method. Developmental research can employ retrospection but should approach such parental reports with circumspection, ensuring that the method is appropriate for the given research question and that adequate efforts have been made to optimize reporter performance. Parental reports are common in basic and applied research and in clinical settings. They are also used routinely outside of academia by teachers and pediatricians (Bornstein & Cote, 2004). The present study cautions individuals who rely on them to attend to the validity and specific nature of retrospective parent reports.

Memory appears to have evolved on functional considerations, and much memory research relates to its adaptive aspects. It should be that parents' memory of children's early health and development has both survival and practical value (Howe & Otgaar, 2013). Understanding how such mnemonic processes influence parenting may position researchers to better understand the complex nature of human caregiving interactions.

Acknowledgements:

We thank Mariya Sumaroka for her contributions to this project. This research was supported by the intramural program of the National Institutes of Health, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development. Kyrsten M. Costlow is now affiliated with the University of Alabama.

References

Allport GW, & Postman L (1947). The psychology of rumor New York, NY: Holt.

Andreano JM, & Cahill L (2009). Sex influences on the neurobiology of learning and memory. Learning and Memory, 16, 248–266. doi: 10.1101/lm.918309 [PubMed: 19318467]

Andreano JM, Dickerson BC, & Barrett LF (2014). Sex differences in the persistence of the amygdala response to negative material. Social Cognitive Affective Neuroscience, 9 1388–1394. doi:10.1093/scan/nst127 [PubMed: 24036962]

Barsky AJ (2002). Forgetting, fabricating, and telescoping: The instability of the medical history. Archives of Internal Medicine, 162, 981–984. doi:10.1001/archinte.162.9.981 [PubMed: 11996606]

Bartlett FC (1932). Remembering, a study in experimental and social psychology. Cambridge, England: Cambridge University Press.

Bates JE, & Bayles K (1984). Objective and subjective components in mothers' perceptions of their children from age 6 months to 3 years. Merrill-Palmer Quarterly, 30, 111–130. Retrieved from http://www.jstor.org/stable/23086228

Bates E, Bretherton I, & Snyder L (1988). From first words to grammar: Individual differences and dissociable mechanisms. Cambridge, England: Cambridge University Press.

Bates E, & Carnevale GF (1993). New directions in research on language development. Developmental Review, 13, 436–470. doi:10.1006/drev.1993.1020

Behar L, & Stringfield S (1974). A behavior rating scale for the preschool child. Developmental Psychology, 10, 601–610. doi:10.1037/h0037058

Belli RF, Shay WL, & Stafford FP (2001). Event history calendars and question list surveys: A direct comparison of interviewing methods. Public Opinion Quarterly, 65, 45–74. Retrieved from http://www.jstor.org/stable/3078785 [PubMed: 11264054]

Bernstein DP, Fink L, Handelsman L, Foote J, Lovejoy M, Wenzel K, … Ruggiero J (1994). Initial reliability and validity of a new retrospective measure of child abuse and neglect. The American Journal of Psychiatry, 151, 1132–1136. doi:10.1176/ajp.151.8.1132 [PubMed: 8037246]

Bisby JA, Horner AJ, Hørlyck LD, & Burgess N (2016). Opposing effects of negative emotion on amygdalar and hippocampal memory for items and associations. Social Cognitive and Affective Neuroscience, 11, 981–990. doi: 10.1093/scan/nsw028 [PubMed: 26969864]

Bluck S (2003). Autobiographical memory: Exploring its functions in everyday life. Memory, 11, 113–123. doi: 10.1080/741938206 [PubMed: 12820825]

Bornstein MH (2014). Parents' reports about their children's lives In Melton GB, Ben-Arieh A, Cashmore J, Goodman GS, & Worley NK (Eds.), The SAGE handbook of child research (pp. 486–533). Los Angeles, CA: Sage.

Bornstein MH (2015). Children's parents In Lerner RM, Bornstein MH, & Leventhal T (Eds.), Handbook of child psychology and developmental science (pp. 1–78). doi:10.1002/9781118963418.childpsy40

Bornstein MH, & Cote LR (2004). "Who is sitting across from me?" Immigrant mothers' knowledge of parenting and children's development. Pediatrics, 114(5), e557–e564. [PubMed: 15520089]

Bornstein MH, Gaughran JM, & Seguí I (1991). Multimethod assessment of infant temperament: Mother questionnaire and mother and observer reports evaluated and compared at five months using the Infant Temperament Measure. International Journal of Behavioral Development, 14, 131–151. doi: 10.1177/016502549101400202

Bornstein MH, Jager J, & Putnick DL (2013). Sampling in developmental science: Situations, shortcomings, solutions, and standards. Developmental Review, 33, 357–370. doi: 10.1016/j.dr.2013.08.003 [PubMed: 25580049]

Bornstein MH, Putnick DL, & Esposito G (2017). Continuity and stability in development. Child Development Perspectives, 11, 113–119. doi:10.1111/cdep.12221 [PubMed: 29503666]

Bornstein MH, Tamis-LeMonda CS, Pascual L, Haynes OM, Painter K, Galperín C, & Pêcheux M-G (1996). Ideas about parenting in Argentina, France, and the United States. International Journal of Behavioral Development, 19, 347–367. doi:10.1177/016502549601900207

Bornstein MH, & Toole M (2010). Assessment of parenting In Tyano S, Keren M, Herrman H, & Cox J (Eds.), Parenthood and mental health: A bridge between infant and adult psychiatry (pp. 349–355). West Sussex, UK: Wiley.

Bower GH (1981). Mood and memory. American Psychologist, 36, 129–148. doi:10.1037/0003-066X.36.2.129 [PubMed: 7224324]

Bowerman M (1985). What shapes children's grammars In Slobin D (Ed.), The crosslinguistic study of language acquisition (pp. 1257–1319). Hillsdale, NJ: Erlbaum.

Brennen T, Hasanović M, Zotović M, Blix I, Solheim Skar AM, Prelić NK, Mehmedović I, Pajević I, Popović N, & Gavrilov-Jerković V (2010). Trauma exposure in childhood impairs the ability to

recall specific autobiographical memories in late adolescence. Journal of Traumatic Stress, 23, 240–247. doi: 10.1002/jts.20513 [PubMed: 20419732]

Brewin CR, Andrews B, & Gotlib IH (1993). Psychopathology and early experience: A reappraisal of retrospective reports. Psychological Bulletin, 113, 82–98. doi:10.1037/0033-2909.113.1.82 [PubMed: 8426875]

Buchanan TW, & Adolphs R (2002). The role of the human amygdala in emotional modulation of long-term declarative memory In Moore S & Oaksford M (Eds.), Emotional cognition: From brain to behavior (pp. 9–34). Amsterdam: John Benjamins Publishing.

Bugental D, & Corpuz R (in press). Parenting attributions In Bornstein MH (Ed.), Handbook of parenting (3rd ed., pp. xx–xx). New York, NY: Routledge.

Burke A, Heuer F, & Reisberg D (1992). Remembering emotional events. Memory & Cognition, 20, 277–290. doi: 10.3758/BF03199665 [PubMed: 1508053]

Cahill L (2003). Sex-related influences on the neurobiology of emotionally influenced memory. Annals of the New York Academy of Sciences, 985, 163–173. doi: 10.1111/j.1749-6632.2003.tb07080.x [PubMed: 12724157]

Canli T, Desmond JE, Zhao Z, & Gabrieli JD (2002). Sex differences in the neural basis of emotional memories. Proceedings of the National Academy of Sciences, 99, 10789–10794. doi: 10.1073/pnas.162356599

Caspi A, Moffitt TE, Thornton A, Freedman D, Amell JW, Harrington H, … Silva PA (1996). The life history calendar: A research and clinical assessment method for collecting retrospective event-history data. International Journal of Methods in Psychiatric Research, 6, 101–114. doi:10.1002/(SICI)1234-988X(199607)6:2<101::AID-MPR156>3.3.CO;2-E

Chess S, Thomas A, & Birch HG (1966). Distortions in developmental reporting made by parents of behaviorally disturbed children. Journal of the American Academy of Child Psychiatry, 5, 226–234. doi:10.1016/S0002-7138(09)62054-9 [PubMed: 5908286]

Chiu YC, Dolcos F, Gonsalves BD, & Cohen NJ (2013). On opposing effects of emotion on contextual or relational memory. Frontiers in Psychology, 4, 103. doi: 10.3389/fpsyg.2013.00103. [PubMed: 23460770]

Christianson S-Å (1992). Emotional stress and eyewitness memory: A critical review. Psychological Bulletin, 112, 284–309. doi: 10.1037/0033-2909.112.2.284 [PubMed: 1454896]

Cohen J (1988). Statistical power analysis for the behavioral sciences (2nd ed.). New York, NY: Routledge.

Conway MA (2009). Episodic memories. Neuropsychologia, 47, 2305–2313. doi: 10.1016/j.neuropsychologia.2009.02.003 [PubMed: 19524094]

Conway MA, Pleydell-Pearce CW (2000). The construction of autobiographical memories in the self-memory system. Psychological Review, 107, 261–288. doi: 10.1037/0033-295X.107.2.261 [PubMed: 10789197]

Cooper C, Dennison EM, Leufkens HG, Bishop N, & van Staa TP (2004) Epidemiology of childhood fractures in Britain: a study using the general practice research database. J Bone and Mineral Research, 19, 1976–1981. doi:10.1359/JBMR.040902

Crowne DP, & Marlowe D (1960). A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 24, 349–354. doi:10.1037/h0047358 [PubMed: 13813058]

D'Alessio M (1990). Social representations of childhood: An implicit theory of childhood In Duveen G & Lloyd B (Eds.), Social representations and the development of knowledge (pp. 70–90). Cambridge, England: Cambridge University Press.

Dallenbach KM (1913). The relation of memory error to time interval. Psychological Review, 20, 323–337. doi: 10.1037/h0076103

Davachi L (2006). Item, context and relational episodic encoding in humans. Current Opinion in Neurobiology, 16, 693–700. doi: 10.1016/j.conb.2006.10.012 [PubMed: 17097284]

Davis PJ (1999). Gender differences in autobiographical memory for childhood emotional experiences. Journal of Personality and Social Psychology, 76, 498–510. doi: 10.1037/0022-3514.76.3.498 [PubMed: 10101879]

DeGrada E, & Ponzo E (1971). La normalita del bambino come pregiudizio dell'adulto [Adult prejudices about infant normality]. Rome, Italy: Bulzoni.

De Los Reyes A, & Kazdin AE (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. Psychological Bulletin, 131, 483–509. doi:10.1037/0033-2909.131.4.483 [PubMed: 16060799]

Digman JM (1990). Personality structure: Emergence of the five-factor model. Annual Review of Psychology, 41, 417–440. doi:10.1146/annurev.ps.41.020190.002221

Dix T (1991). The affective organization of parenting: Adaptive and maladaptative processes. Psychological Bulletin, 110, 3–25. doi:10.1037/0033-2909.110.1.3 [PubMed: 1891517]

Dolcos F, Denkova E, & Dolcos S (2012). Neural correlates of emotional memories: A review of evidence from brain imaging studies. Psychologia, 55, 81–111. doi: 10.2117/psysoc.2012.80

Dolcos F, Katsumi Y, Denkova E, Weymar M, & Dolcos S (2017). Current issues and emerging directions in the impact of emotion on memory: A review of evidence from brain imaging investigations In Tsukiura T, & Umeda S (Eds.), Memory in social context (pp 57–101). Berlin: Springer. doi: 10.1007/978-4-431-56591-8_5

Dolcos F, LaBar KS, & Cabeza R (2005). Remembering one year later: role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. Proceedings of the National Academy of Sciences, 102, 2626–2631. doi:10.1073/pnas.0409848102

Dunn L, & Dunn L (1981). Manual for the Peabody Picture Vocabulary Test-Revised (PPVT-R). Circle Pines, MN: American Guidance Service.

Ebbinghaus H (1913). Memory: A contribution to experimental psychology. New York, NY: Teachers College, Columbia University.

Eichenbaum H (2001). The hippocampus and declarative memory: Cognitive mechanisms and neural codes. Behavioural Brain Research, 127, 199–207. doi: 10.1016/s0166-4328(01)00365-5. [PubMed: 11718892]

Faul F, Erdfelder E, Buchner A, & Lang AG (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods, 41, 1149–1160. doi:10.3758/BRM.41.4.1149 [PubMed: 19897823]

Feldman HM, Dale PS, Campbell TF, Colborn DK, Kurs-Lasky M, Rockette HE, & Paradise JL (2005). Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. Child Development, 76, 856–868. doi:10.1111/j.1467-8624.2005.00882.x [PubMed: 16026501]

Feldman SS, Wentzel KR, & Gehring TM (1989). A comparison of the views of mothers, fathers, and pre-adolescents about family cohesion and power. Journal of Family Psychology, 3, 39–60. doi:10.1037/h0080523

Fenson L, Dale PS, Reznick JS, Bates E, Thal DJ, & Pethick SJ (1994). Variability in early communicative development. Monographs of the Society for Research in Child Development, i–185. doi:10.2307/1166093 [PubMed: 8047076]

Forman DR, O'Hara MW, Larsen K, Coy KC, Gorman LL, & Stuart S (2003). Infant emotionality: Observational methods and the validity of maternal reports. Infancy, 4, 541–565. doi:10.1207/S15327078IN0404_08

Fox J (1997). Applied regression analysis, linear models, and related methods. Thousand Oaks, CA: Sage.

Glass GV, & Hopkins KD (1996). Statistical methods in education and psychology (3rd ed.). Boston, MA: Allyn & Bacon.

Graham S (1992). "Most of the subjects were White and middle class": Trends in published research on African Americans in selected APA journals, 1970–1989. American Psychologist, 47, 629–639. doi: 10.1037/0003-066X.47.5.629

Hamann S (2005). Sex differences in the responses of the human amygdala. Neuroscientist, 11, 288–293. doi: 10.1177/1073858404271981 [PubMed: 16061516]

Hardt J, & Rutter M (2004). Validity of adult retrospective reports of adverse childhood experiences: Review of the evidence. Journal of Child Psychology and Psychiatry, 45, 260–273. doi:10.1111/j.1469-7610.2004.00218.x [PubMed: 14982240]

Henry B, Moffitt TE, Caspi A, Langley J, & Silva PA (1994). On the "remembrance of things past": A longitudinal evaluation of the retrospective method. Psychological Assessment, 6, 92–101. doi:10.1037/1040-3590.6.2.92

Hill LG, & Betz DL (2005). Revisiting the retrospective pretest. American Journal of Evaluation, 26, 501–517. doi:10.1177/1098214005281356

Holland AC, & Kensinger EA (2013). The neural correlates of cognitive reappraisal during emotional autobiographical memory recall. Journal of Cognitive Neuroscience, 25, 87–108. doi: 10.1162/jocn_a_00289 [PubMed: 22905826]

Hollingshead AB (1975). The four-factor index of social status. Unpublished manuscript, Yale University, New Haven, CT.

Howard GS (1980). Response-shift bias. Evaluation Review, 4, 93–106. doi: 10.1177/0193841X8000400105

Howe ML, & Otgaar H (2013). Proximate mechanisms and the development of adaptive memory. Current Directions in Psychological Science, 22, 16–22. doi: 10.1177/0963721412469397

Humes K, Jones NA, & Ramirez RR (2011). Overview of race and Hispanic origin: 2010 (Report No. C2010BR-02). Retrieved from https://www.census.gov/content/dam/Census/library/publications/2011/dec/c2010br-02.pdf

Jackson DN (1994). Manual for the Jackson Personality Inventory-Revised. Port Huron, MI: Sigma Assessment Systems.

Jager J, Bornstein MH, Putnick DL, & Hendricks C (2012). Family members' unique perspectives of the family: Examining their scope, size, and relations to individual adjustment. Journal of Family Psychology, 26, 400–410. doi:10.1037/a0028330 [PubMed: 22545933]

Jager J, Putnick DL, & Bornstein MH (2017). More than just convenient: The scientific merits of homogeneous convenience samples. Monographs of the Society for Research in Child Development, 82, 13–30. doi: 10.1111/mono.12296 [PubMed: 28475254]

Jones IE, Williams SM, Dow N, & Goulding A (2002). How many children remain fracture-free during growth? A longitudinal study of children and adolescents participating in the Dunedin Multidisciplinary Health and Development Study. Osteoporosis International, 13, 990–995. doi:10.1007/s001980200137 [PubMed: 12459942]

Kagan J (1998). Biology and the child In Eisenberg N (Ed) & Damon W (Series Ed.), Handbook of child psychology: Vol. 3. Social, emotional and personality development (pp. 177–235). New York, NY: Wiley.

Kappes A, & Crockett MJ (2016). The benefits and costs of a rose-colored hindsight. Trends in Cognitive Science, 20, 644–646. doi:10.1016/j.tics.2016.06.009

Keenan K, & Wakschlag LS (2002). Can a valid diagnosis of disruptive behavior disorder be made in preschool children? American Journal of Psychiatry, 159, 351–358. doi: 10.1176/appi.ajp.159.3.351 [PubMed: 11869995]

Keenan K, Wakschlag LS, Danis B, Hill C, Humphries M, Duax J, & Donald R (2007). Further evidence of the reliability and validity of DSM-IV ODD and CD in preschool children. Journal of the American Academy of Child & Adolescent Psychiatry, 46, 457–468. doi:10.1097/CHI.0b013e31803062d3 [PubMed: 17420680]

Kensinger EA (2009). Remembering the details: Effects of emotion. Emotion Review, 1, 99–113. doi: 10.1177/1754073908100432 [PubMed: 19421427]

Kouchaki M, & Gino F (2016). Memories of unethical actions become obfuscated over time. Proceedings of the National Academy of Sciences, 113, 6166–6171. doi:10.1073/pnas.1523586113

LaBar KS, & Cabeza R (2006). Cognitive neuroscience of emotional memory. Nature Reviews Neuroscience, 7(1), 54–64. doi:10.1038/nrn1825. [PubMed: 16371950]

Leerkes EM, & Augustine ME (in press). Parenting and emotions In Bornstein MH (Ed.), Handbook of parenting (3rd ed., pp. xx–xx). New York, NY: Routledge.

Lewinsohn PM, & Rosenbaum M (1987). Recall of parental behavior by acute depressives, remitted depressives, and nondepressives. Journal of Personality and Social Psychology, 52, 611–619. doi:10.1037/0022-3514.52.3.611 [PubMed: 3572729]

Lupien S, Gaudreau S, Tchiteya B, Maheu F, Sharma S, … Nair N (1997). Stress-induced declarative memory impairment in healthy elderly subjects: Relationship to cortisol reactivity. The Journal of Clinical Endocrinology and Metabolism, 82, 2070–2075. doi: 10.1210/jc.82.7.2070. [PubMed: 9215274]

MacPhee D, Benson JB, & Bullock D (1986). Influences on maternal self-perceptions. Unpublished manuscript, Colorado State University, Fort Collins, CO.

Macrae CN, Moran JM, Heatherton TF, Banfield JF, & Kelley WM (2004). Medial prefrontal activity predicts memory for self. Cerebral Cortex, 14, 647–654. doi: 10.1093/cercor/bhh025 [PubMed: 15084488]

Mather M (2007). Emotional arousal and memory binding: An object-based framework. Perspectives on Psychological Science, 2, 33–52. doi: 10.1111/j.1745-6916.2007.00028.x [PubMed: 26151918]

Mather M (2016). The affective neuroscience of aging. Annual Review of Psychology, 67, 213–238.doi: 10.1146/annurev-psych-122414-033540

Mather M & Sutherland MR (2011). Arousal-biased competition in perception and memory. Perspectives on Psychological Science, 6, 114–133. doi: 10.1177/1745691611400234 [PubMed: 21660127]

Matlin MW, & Stang DJ (1978). The Pollyanna Principle: Selectivity in language, memory, and thought. Cambridge, MA: Schenkman.

Matt GE, Vázquez C, & Campbell WK (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. Clinical Psychology Review, 12, 227–255. doi:10.1016/0272-7358(92)90116-P

McCloskey M, Wible CG, & Cohen NJ (1988). Is there a special flashbulb-memory mechanism? Journal of Experimental Psychology: General, 117, 171–181. doi: 10.1037/0096-3445.117.2.171

McFarland C, & Buehler R (1998). The impact of negative affect on autobiographical memory: The role of self-focused attention to moods. Journal of Personality and Social Psychology, 75, 1424–1440. doi:10.1037/0022-3514.75.6.1424 [PubMed: 9914662]

Moon RJ, Lim A, Farmer M, Segaran A, Clarke NM, Harvey NC, Cooper C, & Davies JH (2016). Validity of parental recall of children's fracture: Implications for investigation of childhood osteoporosis. Osteoporosis International, 27, 809–813. doi: 10.1007/s00198-015-3287-7 [PubMed: 26286627]

Murray L, Richards MPM, & Nihouarn-Sigurdardottir J (in press). Mothering. In Bornstein MH (Ed.), Handbook of parenting (3rd ed., pp. xx–xx). New York: Routledge.

Paivio SC (2001). Stability of retrospective self-reports of child abuse and neglect before and after therapy for child abuse issues. Child Abuse & Neglect, 25, 1053–1068. doi: 10.1016/S0145-2134(01)00256-3 [PubMed: 11601597]

Parke RD (2000). Beyond White and middle class: Cultural variations in families—assessments, processes, and policies. Journal of Family Psychology, 14, 331–333. doi:10.1037/0893-3200.14.3.331 [PubMed: 11025927]

Parke RD, & Cookston JT (in press). Fathers and families In Bornstein MH (Ed.), Handbook of parenting (3rd ed., pp. xx–xx). New York: Routledge.

Pickett KE, Kasza K, Biesecker G, Wright RJ, & Wakschlag LS (2009). Women who remember, women who do not: A methodological study of maternal recall of smoking in pregnancy. Nicotine & Tobacco Research, 11, 1166–1174. doi:10.1093/ntr/ntp117 [PubMed: 19640836]

Pillemer DB (2003). Directive functions of autobiographical memory: The guiding power of the specific episode. Memory, 11, 193–202. doi:10.1080/741938208 [PubMed: 12820831]

Piolino P, Desgranges B, & Eustache F (2009). Episodic autobiographical memories over the course of time: cognitive, neuropsychological and neuroimaging findings. Neuropsychologia, 47, 2314–2329. doi: 10.1016/j.neuropsychologia.2009.01.020 [PubMed: 19524095]

Pless CE, & Pless IB (1995). How well they remember: The accuracy of parent reports. Archives of Pediatric and Adolescent Medicine, 149, 553–558. doi: 10.1001/archpedi.1995.02170180083016

Pratt CC, McGuigan WM, & Katzev AR (2000). Measuring program outcomes: Using retrospective pretest methodology. American Journal of Evaluation, 21, 341–349. doi: 10.1016/S1098-2140(00)00089-8

Ranganath C (2010). Binding items and contexts the cognitive neuroscience of episodic memory. Current Directions in Psychological Science, 19, 131–137. doi: 10.1177/0963721410368805

Reuben A, Moffitt TE, Caspi A, Belsky DW, Harrington H, Schroeder F… Danese A (2016). Lest we forget: comparing retrospective and prospective assessments of adverse childhood experiences in

the prediction of adult health. Journal of Child Psychology and Psychiatry, 57, 1103–1112. doi: 10.1111/jcpp.12621 [PubMed: 27647050]

Richard W, Skowronski JJ, & Thompson CP (2003). Life is pleasant—and memory helps to keep it that way! Review of General Psychology, 7, 203–210. doi:10.1037/1089-2680.7.2.203

Ritchey M, Dolcos F, & Cabeza R (2008). Role of amygdala connectivity in the persistence of emotional memories over time: An event-related fMRI investigation. Cerebral Cortex 18, 2494–2504. doi:10.1093/cercor/bhm262 [PubMed: 18375529]

Ritchie TD, Skowronski JJ, Wood SE, Walker WR, Vogl RJ, & Gibbons JA (2006). Event self-importance, event rehearsal, and the fading affect bias in autobiographical memory. Self Identity, 5, 172–195. doi:10.1080/15298860600591222

Rubin DC (2006). The basic-systems model of episodic memory. Perspectives on Psychological Science, 1, 277–311. doi: 10.1111/j.1745-6916.2006.00017.x [PubMed: 26151796]

Reynell JK, & Gruber CP (1990). Reynell Developmental Language Scales U. S. Edition. Los Angeles, CA: Western Psychological Services.

Reynolds WM (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. Journal of Clinical Psychology, 38, 119–125. doi:10.1002/1097-4679(198201)38:1<119::AID-JCLP2270380118>3.0.CO;2-I

Robbins LC (1963). The accuracy of parental recall of aspects of child development and of child rearing practices. The Journal of Abnormal and Social Psychology, 66, 261–270. doi:10.1037/h0049084. [PubMed: 13974231]

Ros L, & Latorre JM (2010). Gender and age differences in the recall of affective autobiographical memories using the autobiographical memory test. Personality and Individual Differences, 49, 950–954. doi: 10.1016/j.paid.2010.08.002

Rutter SM, Maughan B, Pickles A, & Simonoff E (1998). Retrospective recall recalled In Cairns RB, Bergman LR, & Kagan JE (Eds.), Methods and models for studying the individual (pp. 219–243). Thousand Oaks, CA: Sage.

Sakaki M, Niki K, & Mather M (2012). Beyond arousal and valence: the importance of the biological versus social relevance of emotional stimuli. Cognition Affect. Behavioral Neuroscience 12, 115–139. doi: 10.3758/s13415-011-0062-x.

Sameroff AJ, Seifer R, & Elias PK (1982). Sociocultural variability in infant temperament ratings. Child Development, 53, 164–173. Retrieved from http://www.jstor.org/stable/1129649 [PubMed: 7060419]

Saywitz KJ, Geiselman RE, & Bornstein GK (1992). Effects of cognitive interviewing and practice on children's recall performance. Journal of Applied Psychology, 77, 744–756. doi:10.1037/0021-9010.77.5.744

Schacter DL (2007). How the mind forgets and remembers: The seven sins of memory. New York, NY: Souvenir Press.

Schacter DL, Koutstaal W, & Norman KA (1997). False memories and aging. Trends in Cognitive Sciences, 1, 229–236. doi: 10.1016/S1364-6613(97)01068-1 [PubMed: 21223912]

Schwarz N (1999). Self-reports: How the questions shape the answers. American Psychologist, 54, 93–105. doi:10.1037/0003-066X.54.2.93

Schwarz N, & Oyserman D (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. American Journal of Evaluation, 22, 127–160. doi:10.1177/109821400102200202

Seidlitz L, & Diener E (1998). Sex differences in the recall of affective experiences. Journal of Personality and Social Psychology, 74, 262–271. doi: 10.1037/0022-3514.74.1.262 [PubMed: 9457787]

Seifer R, Sameroff AJ, Barrett LC, & Krafchuck E (1994). Infant temperament measured by multiple observations and mother report. Child Development, 65, 1478–1490. doi:10.1111/j.1467-8624.1994.tb00830.x [PubMed: 7982363]

Sparrow SS, Balla DA, & Cicchetti DV (1984). Vineland Adaptive Behavior Scales: Interview Edition, Survey Form Manual. Circle Pines, MN: American Guidance Service.

Stern W (1902). Zur psychologie der aussage. Berlin, Germany: J. Gutentag.

Stevens JS, & Hamann S (2012). Sex differences in brain activation to emotional stimuli: a meta-analysis of neuroimaging studies. Neuropsychologia, 50, 1578–1593. doi: 10.1016/j.neuropsychologia.2012.03.011 [PubMed: 22450197]

Sutterer DW, & Edward A (2016). Retrieval practice enhances the accessibility but not the quality of memory. Psychonomic Bulletin & Review, 23, 831–841. doi:10.3758/s13423-015-0937-x [PubMed: 26404635]

Thal D, & Bates E (1990). Continuity and variation in early language development In Colombo J & Fagen J (Eds.), Individual differences in infancy: Reliability, stability, prediction (pp. 359–381). Hillsdale, NJ: Erlbaum.

Tilton-Weaver LC, & Kakihara F (2008). United States of America In Arnett JJ (Ed.), International encyclopedia of adolescence (pp. 1061–1076). New York, NY: Routledge.

Tomlinson M, Bornstein MH, Marlow M, & Swartz L (2014). Imbalances in the knowledge about infant mental health in rich and poor countries: Too little progress in bridging the gap. Infant Mental Health Journal, 35, 624–629. doi:10.1002/imhj.10078 [PubMed: 25798511]

Tulving E (2005). Episodic memory and autonoesis: Uniquely human? In Terrace HS, & Metcalfe J (Eds.), The missing link in cognition (pp. 3–56). New York, NY: Oxford University Press.

Unsworth N (2010). On the division of working memory and long-term memory and their relation to intelligence: A latent variable approach. Acta Psychologica, 134, 16–28. doi: 10.1016/j.actpsy.2009.11.010 [PubMed: 20022311]

Unsworth N, Brewer GA, & Spillers GJ (2011). Variation in working memory capacity and episodic memory: Examining the importance of encoding specificity. Psychonomic Bulletin & Review, 18, 1113–1118. doi: 10.3758/s13423-011-0165-y [PubMed: 21912997]

Vaughn BE, Bradley CF, Joffe LS, Seifer R, & Barglow P (1987). Maternal characteristics measured prenatally are predictive of ratings of temperamental “difficulty” on the Carey Infant Temperament Questionnaire. Developmental Psychology, 23, 152–161. doi:10.1037/0012-1649.23.1.152

von Eye A (2002). Configural frequency analysis: Methods, models, and applications. Mahwah, NJ: Erlbaum.

Walker WR, & Skowronski JJ (2009). The fading affect bias: But what the hell is it for? Applied Cognitive Psychology, 23, 1122–1136. doi: 10.1002/acp.1614

Weymar M, & Hamm AO (2013). Electrophysiological signature of emotional memories In Linden M, & Rutkowski K (Eds.), Hurting memories and beneficial forgetting: Posttraumatic stress disorders, biographical developments and social conflicts (pp. 21–35). Amsterdam: Elsevier. doi: 10.1016/B978-0-12-398393-0.00002-X

Whalley MG, Rugg MD, & Brewin CR (2012). Autobiographical memory in depression: An fMRI study. Psychiatry Research: Neuroimaging, 201, 98–106. doi: 10.1016/j.pscychresns.2011.08.008

Widom CS, & Morris S (1997). Accuracy of adult recollections of childhood victimization, part 2: Childhood sexual abuse. Psychological Assessment, 9, 34–46. doi:10.1037/1040-3590.9.1.34.

Williams LM (1994). Recall of childhood trauma: A prospective study of women’s memories of child sexual abuse. Journal of Consulting and Clinical Psychology, 62,1167–1176. doi: 10.1037/0022-006X.62.6.1167 [PubMed: 7860814]

Wixted JT, & Wells GL (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. Psychological Science in the Public Interest, 18, 10–65. doi: 10.1177/1529100616686966 [PubMed: 28395650]

Yaoi K, Osaka M, & Osaka N (2015). Neural correlates of the self-reference effect: evidence from evaluation and recognition processes. Frontiers in Human Neuroscience, 9, 383. doi: 10.3389/fnhum.2015.00383 [PubMed: 26167149]

Yarrow MR, Campbell JD, & Burton RV (1970). Recollections of childhood: A study of the retrospective method. Monographs of the Society for Research in Child Development, 35 (Serial 138), iii–83. doi: 10.2307/1165649

Yonelinas AP, & Ritchey M (2015). The slow forgetting of emotional episodic memories: an emotional binding account. Trends in Cognitive Sciences, 19, 259–267. doi: 10.1016/j.tics.2015.02.009 [PubMed: 25836045]

Appl Dev Sci. Author manuscript; available in PMC 2021 January 01.

Table 1

Domain 1: Consistency in means and relative standing of original and retrospective maternal reports of mother cognitions

	Original			Retrospective				
Measure	**α**	***M***	***SD***	**α**	***M***	***SD***	***t***	***r***
5 months							**2.19**[a]	**.37**[a]
Satisfaction with parenthood-employment balance	--	3.63	.98	--	3.53	1.20	−.59	.56***
Satisfaction with childcare arrangements	--	4.46	.96	--	4.24	.80	−1.16	.18
Support from the child's father	--	3.70	.63	--	4.40	.94	4.76***	.30*
Support from my family and relatives (mother's mother)	--	2.19	1.65	--	2.86	1.49	2.56*	.40**
Support from the family and relatives of the child's father (mother-in-law)	--	1.43	1.50	--	2.43	1.68	5.20***	.68***
Support from child's pediatrician	--	2.84	1.00	--	3.67	.97	4.60***	.26
Support from friends or neighbors	--	1.89	1.32	--	2.89	1.63	3.85***	.33*
Support from an organized course, class, or group	--	.40	.95	--	.63	1.33	1.26	.48***
Support from books, magazines, newspapers	--	3.18	1.06	--	3.00	1.46	−.76	.24
20 months							**−.19**[a]	**.34**[a]
SPPR Investment	.74	3.02	.79	.72	3.17	.75	1.10	.36*
SPPR Competence	.74	3.98	.58	.86	3.90	.80	−.74	.56***
SPPR Role Balance	.66	3.70	.66	.79	3.85	.78	1.14	.26
SPPR Satisfaction	.61	4.72	.33	.86	4.52	.64	−2.26*	.39**
48 months							**−1.96**[a]	**.29**[a]
Satisfaction with parenthood-employment balance	--	3.76	.82	--	3.82	1.09	.40	.63***
Satisfaction with childcare arrangements	--	4.62	.56	--	4.24	.79	−2.37*	.22
Mother's perceived influence on the child	.89	5.35	.54	.91	4.85	.71	−3.92***	.20

Note: -- = single item, or Cronbach alpha not appropriate.

[a]Grand index.

*$p \leq .05$.

**$p \leq .01$.

***$p \leq .001$.

Table 2

Domain 2: Consistency in means and relative standing of original and retrospective maternal reports of mother-reported child and mother behaviors

	Original			Retrospective				
Measure	**α**	***M/ % yes***	***SD***	**α**	***M/ % yes***	***SD***	***t/ χ^2***	***r /% consistency***
Maternal Report of Child Behaviors								
5 months							**−.01** [a]	**.66** [a]
Birth weight (g)	--	3608.98	477.05	--	3603.75	478.79	−.54	.99***
Birth length (cm)	--	52.43	2.53	--	52.55	3.97	.15/.64[b]	−.07/.46** [b]
Maturity at birth (% term)	--	100%		--	95%		n.a.	95%
Ease of “settling in” during 1st month at home	--	4.07	1.20	--	4.15	.95	.51	.52***
Absence of difficult behaviors during 1st month at home	--	9.32	.91	--	9.30	1.02	−.13	.32*
20 months							**2.43** [a]	**.21** [a]
VABS communication	--	1.20	.28	--	1.50	.27	6.33***	.39*
VABS daily living skills	--	1.29	.31	--	1.26	.89	−.20	.10
VABS socialization	--	.78	.40	--	1.02	.49	2.67**	.19
VABS motor	--	1.19	.39	--	1.26	.48	.90	.28
48 months							**2.64** [a]	**.49** [a]
PBQ non-hostile	.70	16.29	2.52	.80	18.29	2.84	5.44***	.65***
PBQ non-anxious	.67	14.05	2.16	.76	15.45	2.36	4.15***	.58***
PBQ non-hyper	.81	4.69	1.69	.79	5.38	2.18	2.39*	.59***
PBQ total	.82	44.55	5.42	.87	49.49	6.28	6.58***	.64***
VABS communication	--	1.52	.33	--	1.54	.40	.42	.64***
VABS daily living skills	--	1.16	.32	--	1.34	.50	2.23*	.35*
VABS socialization	--	1.27	.40	--	1.51	.40	3.13**	.28
VABS motor	--	1.52	.31	--	1.75	.26	4.73***	.44**
Child enjoyment	.64	5.04	.42	.84	4.69	.63	−3.19**	.30
Expectations of child performance	.82	4.38	.65	.91	4.45	.82	.54	.47**

	Original			Retrospective				
Measure	**α**	***M/ % yes***	***SD***	**α**	***M/ % yes***	***SD***	***t/ χ^2***	***r /% consistency***
Maternal Report of Mother Behaviors								
5 months							**9.35** [a]	**81.25%** [a]
Planned pregnancy	--	60%		--	86%		3.89 *	69%
Problems or complications during the pregnancy	--	26%		--	24%		16.11 ***	87%
Problems or complications during labor and delivery	--	38%		--	51%		8.05 **	72%
Prenatal care	--	100%		--	97%		n.a.	97%
Hours of employment	--	31.76	15.18	--	30.57	16.79	−.79	.81 ***
20 months							**.18** [a]	**.55** [a]
PSQ Social	.69	4.45	.36	.80	4.34	.42	−2.19 *	.71 ***
PSQ Didactic	.51	3.79	.34	.69	3.80	.43	.30	.49 ***
PSQ Limit Setting	.78	3.89	.80	.72	3.98	.71	.79	.60 ***
Hours of employment	--	24.90	18.37	--	29.00	16.81	1.80	.65 ***
48 months								
Hours of employment	--	26.05	18.66	--	25.25	18.81	−.42	.79 ***

Note. n.a.= statistic was not available because the chi-square could not be computed due to missing cells. -- = single item, or Cronbach alpha not appropriate.

[a] Grand index.

[b] Values after the slash were after removing two influential outliers.

* $p \leq .05$.

** $p \leq .01$.

*** $p \leq .001$.

Table 3

Domain 3: Consistency in means and relative standing of original and retrospective maternal reports of observed child and mother behaviors

	Original		Retrospective			
Measure	***M***	***SD***	***M***	***SD***	***t***	***r***
Observed Behavior of Child						
5 months					**4.63** [a]	**.01** [a]
Babbling or cooing	--	--	--	--	n.a.	−.37 *
Fussing or crying	--	--	--	--	n.a.	−.24
Highest developmental level of sitting	.36	.22	.53	.25	3.68 ***	.24
Highest developmental level of moving	.33	.22	.54	.24	5.57 ***	.42 **
20 months					**1.93** [a]	**.12** [a]
RDLS Pointing to objects	6.51	1.81	6.44	1.64	−.20	.10
RDLS Following 2-step directions	1.08	1.14	2.68	1.24	6.68 ***	.20
RDLS Developmental level of speech	5.31	1.20	5.43	.99	.55	.19
RDLS Spoken words when presented with the object	3.63	2.31	3.95	2.14	.66	.06
Observed Behavior of Mother						
5 months						**−.10** [a]
Talking	--	--	--	--	n.a.	−.05
Social Play	--	--	--	--	n.a.	−.12
Encourage attention to mother	--	--	--	--	n.a.	−.29
Encourage attention to objects	--	--	--	--	n.a.	−.10
Overall style of directing your baby's attention	--	--	--	--	n.a.	−.03
20 months						**.04** [a]
Play demonstrations	--	--	--	--	n.a.	−.07
Play solicitations	--	--	--	--	n.a.	.10
Social play	--	--	--	--	n.a.	.11

Note. n.a.= statistic was not available because the means were not comparable.

[a]Grand index.

Author Manuscript

Author Manuscript

Author Manuscript

* $p \leq .05$.

** $p \leq .01$.

*** $p \leq .001$.

Table 4

Mothers' confidence in their memories

Domain	*M*	*SD*	Range
Child			
5 months	3.84	1.59	1-7
20 months	4.07	1.30	2-7
48 months	4.84	1.06	2-6
Mother			
5 months	4.60	1.60	2-7
20 months	4.58	1.29	1-6
48 months	5.13	1.06	2-7
Employment	6.36	.78	4-7
Childcare	5.98	1.42	1-7
Social support	5.98	.97	3-7

Note. Possible range = 0-7.