Do More Hours in Center-based Care Cause More Externalizing Problems? A Cross-National Replication Study

Catalina Rey-Guerra*[a], Henrik Daae Zachrisson[b,a], Eric Dearing[a,b], Daniel Berry[c], Susanne Kuger[d,e], Margaret R. Burchinal[f], Ane Nærde[g], Thomas van Huizen[h] & Sylvana Coté [i,j]

[a] Boston College, Chestnut Hill, MA, USA; [b] University of Oslo, Oslo, Norway; [c] University of Minnesota, Minneapolis, MN, USA; [d] German Youth Institute, Munich, Germany; [e] German Institute for International Educational Research, Frankfurt am Main, Germany; [f] University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [g] The Norwegian Center for Child Behavioral Development, Oslo, Norway; [h] Utrecht University, Utrecht, Netherlands; [i] University of Montreal, Montreal, Canada; [j] University of Bordeaux, Bordeaux, France.

*Corresponding author: Catalina Rey-Guerra, Campion Hall Room 201A, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, reyguerr@bc.edu

How to cite

Rey-Guerra, C., Zachrisson, H. D., Dearing, E., Berry, D., Kuger, S., Burchinal, M. R., Nærde, A., van Huizen, T., & Côté, S. M. (2022). Do more hours in center-based care cause more externalizing problems? A cross-national replication study. *Child Development*, 00, 1– 20. https://doi.org/10.1111/cdev.13871

Abstract

Whether high quantities of center-based child care cause behavior problems is a controversial question. Studies using covariate adjustment for selection factors have detected relations between center care and behavior problems, but studies with stronger internal validity less often find such evidence. We examined whether within-child changes in hours in center-based care predicted changes in externalizing problems in toddlers and preschoolers (N = 10,105; 49% female; data collection 1993 to 2012) in seven studies, including from Germany, Netherlands, Norway, two from Canada and two from the U.S. Race/ethnicity data were only collected in the U.S. (57% and 80% White; 42% and 13% African-American; 1.2% and 5% Latinx). Meta-analyses showed no association (r=.00, p=.88) between hours in center-based care and externalizing problems.

Do More Hours in Center-based Care Cause More Externalizing Problems? A Cross-National Replication Study

The question of whether time spent in child care can lead to elevated behavior problems remains a controversial issue. In part, controversy stems from mixed research findings. Some studies indicate that large amounts of time in child care, particularly center-based care, may pose a risk for developing problems, namely externalizing behavior problems such as aggression (Huston et al., 2015; NICHD Early Child Care Research Network [NICHD ECCRN], 2003; Belsky et al., 2007). Other studies, however, find no risk associated with time in child care (e.g., Zachrisson et al., 2013), and some even indicate nonparental care may decrease problem behaviors, especially among children of socially disadvantaged families (Côté et al., 2007; Crosby et al., 2010; Orri et al., 2019).

The fact that child care studies are predominantly correlational rather than experimental complicates efforts to understand the mixed findings. The internal validity of studies examining associations between amount of time in child care (i.e., any type of nonparental care including center-based) and behavior problems has been called into question (Dearing et al., 2015; McCartney et al., 2010). In addition, the "treatment" and counterfactual conditions have varied across studies (e.g., time in maternal care, time in parental care, and time in care other than centers have all been used as counterfactuals), making it challenging to compare results (Dearing & Zachrisson, 2017). Furthermore, there is a relative shortage of research, with strong internal validity, on this topic from outside the United States, which brings up questions of generalizability (Dearing & Zachrisson, 2017; Duncan et al., 2014).

The present study was designed with these limitations to existing research in mind. To extend the cumulative knowledge on this topic, we bring together data from seven prospective longitudinal studies that were conducted in five countries (i.e., Canada, Germany, Netherlands, Norway, and the United States). With these data, we examine associations

between hours in center-based child care and externalizing behavior problems. We focused on time in center care as the "treatment" given evidence that extensive time in this type of care is the most likely to cause problem behaviors (e.g., NICHD ECCRN, 2003; Belsky et al., 2007). To help address internal validity concerns, our primary analyses examine within-child associations between weekly hours in center-based care and behavior problems, allowing us to rule out unmeasured time-invariant potential sources of bias.

A Brief History of the Study of Child Care and Externalizing Problems

As women's participation in the labor force rapidly grew in economically developed countries during the last decades of the 20th century, so too did the proportion of children in nonparental care (Organisation for Economic Cooperation Development [OECD], 2020). Observing this trend led to concerns among some researchers about potential harms of nonparental care, in particular of nonmaternal care, on children's development.

Initial empirical findings indicated that attending nonmaternal child care at an early age and for extensive periods of time (more than 20 hours per week) could heighten the risk of behavioral problems (e.g., Belsky, 1986). In these early studies, heavily influenced by attachment theory and sociocultural attitudes that situated mothers as the primary caregivers of children, any form of nonmaternal care (including paternal care) was the "treatment" or potential risk factor and time in maternal care was the counterfactual condition. Other researchers, however, questioned these early findings arguing that the research was based on non-experimental studies that had not adequately controlled for potential confounds such as quality of care (e.g., Clarke-Stewart, 1989; McCartney & Rosenthal, 1991).

Motivated by these early findings and critiques, the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) launched the Study of Early Child Care and Youth Development (SECCYD) in the early 1990s. The primary focus of this prospective longitudinal study was to examine the effects of time in *nonmaternal* child care

on the behavioral development of a sample of children born in ten locations in the United States. Although non-experimental and not representative of the U.S. population, the SECCYD was novel in its extensive assessments of child care use and developmental outcomes as well as rich family data that helped contextualize child care experiences. With this in mind, results from the SECCYD have strongly impacted the field's understanding of the role of child care in development (e.g., Jacob, 2009). Regarding the effects of quantity of child care on behavior problems, however, the story from the SECCYD has been less than straightforward, and the "treatment" and counterfactual conditions of interest have evolved.

While initial results from the SECCYD indicated negligible risk associated with time in nonmaternal care for behavior problems in early childhood (NICHD ECCRN, 1998), a second wave of reports indicated high levels of hours in non*parental* care (controlling for quality, time spent in center-based care, peer-group exposure, and instability of care) were associated with increased risk of externalizing behavior problems at ages 2 and 4 ½ years but not at 3 years. This second wave of reports also indicated that risk was particularly evident for children in center-based care and when examining caregiver/teacher reports of behavior rather than parent reports (NICHD ECCRN, 2003, 2004). Subsequent SECCYD studies were consistent in finding that teachers reported more behavior problems for children who attended center-based care up to sixth grade (Belsky, et al., 2007). In these studies that highlighted the risk of center-based care, the analyses involved estimating the proportion of time spent in center-based care; the counterfactual was proportion of time in all other forms of care, parental or otherwise (e.g., family child care). Other studies in the U.S. obtained results that were consistent with the SECCYD. For example, using the Early Childhood Longitudinal Study–Birth Cohort (ECLS–B), Coley et al., (2013) found negative associations between time in center care at age 4 and early externalizing behaviors both for parent and

teacher reports. They found that center-based care was a risk regardless of whether (a) parental care or (b) other forms of nonparental care were treated as the counterfactual.

In sum, multiple studies in the U.S. report associations supporting the possibility that long hours in center care poses a risk for externalizing problems compared with both parental care and other forms of nonparental care. There is good reason, however, to be concerned with the internal validity of much of this evidence base (Dearing & Zachrisson, 2017). Nearly all of these studies have used covariate adjustment for observable selection factors, a technique limited by the ubiquitous possibility of unobserved sources of bias. Few studies use more rigorous approaches for ruling out both observed and unobserved sources of bias.

Concerns over Internal Validity

A notable change in the analytical approach to this research question came with McCartney and colleagues' (2010) re-analyses of the SECCYD data, suggesting that the evidence was more mixed than captured by earlier studies. These re-analyses increased analytic rigor aimed at internal validity, with the research team using multiple methods to attempt to control for potential selection bias (i.e., the possibility that children who are more likely to exhibit problem behaviors are also more likely to be in large amounts of center care). The authors found limited and "equivocal" support for the causal hypothesis that extensive time in care leads to more problems (e.g., McCartney et al., 2010, p. 1).

Given that most studies, particularly those pre-dating McCartney et al. (2010), linking extensive time in center-based care with externalizing behavior problems are exclusively correlational, there are serious concerns over potential selection effects. Stated simply, without random assignment, the concern is that unmeasured child, family, and context factors may be the true cause of observed behavioral differences between children in more or less care, rather than time in child care itself.

With the aim of reducing potential bias caused by selection effects, most studies have relied on large sets of statistical controls (e.g., maternal education, family income, maternal employment status, family structure indicators, and parent psychological well-being) to estimate the effects of children being exposed to varying amounts of care. Regarding family income and socioeconomic status, evidence suggests that the association between hours spent in center care and behavior problems, is most evident among families with high incomes or low risk of facing adversities (Berry, 2014; Huston et al., 2015). The main limit of this approach is that the estimates are only valid if all the correct selection factors have, in fact, been controlled, an issue that cannot be empirically known. Even with extensive covariate sets, the potential for unobserved confounders remains. With studies of child care, for example, there is concern for genetic factors, difficult to measure aspects of human capital (e.g., knowledge and experience not gained via formal education), and social capital not captured by traditional family structure or social support measures.

In a review of studies using research designs reducing or removing unobserved confounders, the evidence for such an association was mixed, at best (Dearing & Zachrisson, 2017). Even within studies, associations between time in child care and behavior problems that are evident using covariate adjustment can disappear (or indicate that time in care *reduces* problems) when analytic techniques with stronger internal validity (e.g., instrumental variables and fixed effects estimators) are employed (Crosby et al., 2010; McCartney et al., 2010). These findings raise concerns about whether correlational studies have drawn erroneous conclusions about effects of child care quantity on externalizing problems.

Concerns over External Validity: International Studies

A second concern with the current literature is one of external validity, and especially whether findings from the U.S. can be generalized to other sociopolitical contexts (e.g., Huston et al., 2015; Dearing & Zachrisson, 2017). When considering studies disentangling

center-based care (from other forms of nonparental care) outside the U.S., there is some evidence, albeit mixed, that center-based care poses a developmental risk. Using data from Australia, for example, Yamauchi and Leigh (2011) found that children who attended full-time center-based care exhibited higher reactivity scores than children exclusively in parental care; children in other forms of nonparental care did not significantly differ from those in parental care. In Switzerland, a study examining "group-based" care arrangements (analogous to center care and distinguished from "individual" care settings that included family, neighbor, and day-care homes), found a positive link between time spent in group-based care and externalizing behaviors compared with parental care, but this difference was not evident for other forms of nonparental child care versus parental care (Averdijk, et al., 2011). In England, results have been mixed with Melhuish (2010) reporting negative associations between center-based care and externalizing problems in early childhood (with this association disappearing by age 10), but Barnes and colleagues (Barnes, et al., 2010) finding no evidence of risk associated with center-based care during the first three years compared with other forms of care, parental or non-parental. And, in Canada, Borge and colleagues (2004) found that *more aggressive behaviors* were exhibited by children in exclusive maternal care than those in attending group day-care, a pattern replicated by Côté and colleagues (2007) for children of mothers with low education.

However, each of these studies of center-based care from Australia, Switzerland, England, and Canada relied on observed confounders to address potential selection effects. Among the international studies using strong causal identification strategies (e.g., natural experiments), there is some evidence that center care is a risk. In a German natural experiment, for example, Felfe and Zierow (2018) exploit a reform that increased the proportion of full-day versus half-day slots for public preschools. These authors found negative effects on child behavior compared with a counterfactual of half-day experiences in

other forms of care (parental and/or nonparental). Similarly, in a Canadian natural experiment (e.g., Kottelenberg & Lehrer, 2017) generally adverse effects on behavior were found for a state-wide scale up of child care centers; here, the "treatment" was increasing availability of child care center slots and the counterfactual was the lack thereof. However, in a Norwegian study of exogenous variations in amount of time spent in center care (determined by child birth date and age of entry restrictions) negligible effects on externalizing behavior were documented (Dearing et al., 2015).

In each of these more rigorous international studies, publicly-funded centers were the "treatment" of interest, and all other forms of care (i.e., parental or otherwise) provided the counterfactual, but amount of time in care was estimated quite differently (i.e., part-time versus full-time, any center care versus no center care, and number of months in full-time center care). These diverging results, notably addressing related but different research questions, point to the potential value of comparable studies using the same designs and asking identical research questions across sociopolitical contexts.

Theoretical Explanations for Risk of Center Care

While societal trends rather than theory were the primary initial drivers of much of the work on the topic in the United States and internationally, theoretical arguments have been proposed for why time in center-based care may be a risk for behavioral development. Prominent early explanations focused on disruptions of parent-child attachment, either due directly to the separation anxiety experienced by the child or due more indirectly to the way time apart could undermine parents' abilities to sensitively respond to children's attachment needs (see McCartney & Rosenthal, 1991, for review). However, little evidence has been produced that is consistent with either of these explanations (e.g., Zachrisson et al., 2020); indeed, the exceptional risk of center care versus all other forms of nonparental care suggests that disrupted parent-child attachment is not an explanatory mechanism.

An alternative explanation was based on social learning theory. According to this perspective, large amounts of center-based care could provide children with more opportunities to observe deviant behavior in their peers, which could then be modelled (Clarke-Stewart, 1989). McCartney et al. (2010), for example, found that more time with large groups of peers while in child care predicted a greater likelihood of children displaying problem behaviors. Another study found that, conditional on hours in care, within-child increases in exposure to peer problem behavior in center-based care was associated with within-child increases in aggressive behavior (Ribeiro & Zachrisson, 2019).

Finally, some have hypothesized that quantity of child care may only be harmful or may be most harmful if the quality of care is low. Indeed, quality of center care defined in terms of both structure (e.g., student-teacher ratio) and process (e.g., responsive interactions) may play distinctive roles and challenges for children's adaptation to the new environment. For instance, more experienced caregivers might provide more sensitive interactions to children, while larger classrooms with more children might hinder the amount of interaction opportunities with the caregiver. And, there is evidence to support this hypothesis that quality of care can moderate quantity of child care effects (McCartney et al., 2010).

Yet, with regard to the possible exceptional risk of center-based care, the quality hypothesis then evokes the question of whether center-based care is exceptionally likely to be lower in quality compared to other nonparental care. Evidence from the SECCYD indicates that larger child-staff ratios, which were more often present in centers, predicted lower process quality (i.e., sensitive and stimulating caregiver interactions with children); and, process quality was, on average, lower in centers in that study (e.g., Vandell, 1996; NICHD ECCRN, 2002a). Nonetheless, the overall evidence as to whether there is a behavioral risk, by any mechanism, has remained mixed (Dearing & Zachrisson, 2017), and there is reason to

suspect that issues of both internal and external validity may contribute to inconsistent findings. The present study was designed to help address these issues.

The Present Study

The present confirmatory study extends the cumulative knowledge on associations between hours in center-based care and externalizing behavior problems in three main ways. First, this is the first study on the topic to bring together analyses of data from multiple countries with diverse sociopolitical contexts, allowing us to address concerns of generalizability and replicability in the literature. Specifically, we analyzed longitudinal data from two Canadian and two U.S. studies as well as studies from Germany, the Netherlands, and Norway. Importantly, these countries vary with regard to relevant social policies for family leave and public provision and regulation of Early Childhood Education and Care (ECEC). For example, enrollment rates vary from below 20% (for 0-2 year-olds in Germany) to over 90% (for 3-5 year-olds in Norway), and maternity leave varies from 12 weeks of unpaid leave in the U.S. to 53 weeks with almost full pay in Norway. In Table 1, we have provided further details on country economic, family leave policy, and ECEC indicators, including general information about the countries' child care landscape in terms of quality and ECEC regulations at the time when the studies were taking place.

Table 1

Key Country Indicators for the Five Countries in the Study

Country and years of data collection of the study	GDP per capita[a]	Gini coefficient (income after taxes and transfers)	Government expenditure on pre-primary education as % of GDP (%)	Type of parental leave policy 2004-2005	Enrollment in licensed child care[d] (%)	Pupil-teacher ratio in pre-primary education[f]	Public funding for access to ECEC	ECEC quality regulations
Canada **1999-2002** **2005-2007**	$37,432 in 2000 $41,647 in 2007	33.2 in 1998 33.8 in 2007	0.22 in 1998 0.21 in 2005	• Maternity leave: 15 weeks • Paternal leave: 10 days • 55% of earnings	Children 0-2 years: 24.0 in 2005[e] Children 3-5 years: 24.0 in 2005[e]	17.72 in 2000	Early Childhood Development (ECD) Agreement (2000) provided the provinces/territories with federal funds to improve and expand services and supports for children under age 6 and their families. The Multilateral Framework on Early Learning and Child Care (2003) transferred federal funds to the provinces/territories to improve access to affordable, quality regulated child care.	In center based child care services, 2 out of every 3 educators are required to have formal training, i.e. app 14 years of education of which the last two years are focused on early childhood education. A minimal training of 30 hours is required for educators in family based settings.
Germany **2005-2007**	$40,473 in 2007	31.8 in 2006 31.1 in 2008	0.34 in 1998 0.53 in 2005	• Maternity leave: 6 weeks before birth and 8 weeks after • 100% average annual earnings	Children 0-2 years: 16.8 in 2005 26.8 in 2010 Children 3-5 years: 87.6 in 2005 94.2 in 2010	17.51 in 1993 8.17 in 2014	Universal access was guaranteed for all children in Germany age 3 and older (up until school entry).	Starting in 2002 the 16 federal states introduced curricular guidelines that were mandatory in some states and of more advisory character in others (latest release in 2011). All states demanded minimum standards for teacher-child ratio (ranging between 8 and 11 children age 3-6 years per

								teacher) and minimum teacher qualification (either at least one or two teachers with a five-year tertiary vocational training per group of about 20-25 children).
Netherlands 2010-2012	$45,524 in 2010	27.8 in 2010 27.6 in 2012	0.4 in 2010 0.4 in 2012	• Maternity leave: 16 weeks • 100% of earnings • Additional unpaid parental leave of 6 months	Children 0-2 years: 16.8 in 2005 26.8 in 2010 Children 3-5 years: 94.3 in 2010 94.1 in 2012	15.50 in 2014	Childcare Act (2005) and the Welfare Act included prekindergarten for two- to three-year-olds, day care centers operating in the private market and families received day care subsidies and were available from age 0. ECEC services included center-based daycare, family daycare, and (targeted) preschools. Dual-earner families were eligible for income dependent child care subsidies (through the fiscal system) when using center-based and family daycare. Preschools were partially funded through municipalities and partially funded through low family fees.	Quality requirements included child:staff ratios (varying from 8:1 for 3-year-olds to 4:1 for 0-year-olds). ECEC staff was required to have a vocational degree in ECEC.
Norway 2006-2008	$65,083 in 2007	26.5 in 2006 27 in 2008	0.56 in 1998 0.71 in 2012	• Maternal leave: 43 weeks at 100% of earnings or 53 weeks at 80% of earnings • Paternal leave: 5 weeks	Children 0-2 years: 32.7 in 2005 52.6 in 2010 Children 3-5 years: 87.7 in 2005 96.2 in 2010	6.21 in 1998	Universal ECEC from age one through school entry was publicly funded. Parents paid a fee of app. USD 300/month (with lower prices for low income parents). All ECEC centers were required to follow quality requirements.	Quality requirements included child:staff ratio (3:1 < age 3, 6:1 >= age 3), and one ECEC teacher with a tertiary degree in early childhood education per 9 children < 3 years, and 18 children >= 3 years. A national framework plan (general curriculum) direct the content of ECEC pedagogy.

United States 1993-1996 2003-2004	$37,062 in 1990 $47,158 in 2003	38.2 in 1991 40.5 in 2004 41.4 in 2007	0.33 in 1998 0.34 in 2012	• Parental leave: 12 weeks of unpaid leave * Five States paid temporary disability benefit for 10 weeks.	Children 0-2 years: 27.4 in 2006 25.8 in 2010 Children 3-5 years: 66.3 in 2006 65.8 in 2010	27.23 in 1995 13.82 in 2014	Targeted ECEC for low-income children through federal Early Head Start and Head Start programs, which expanded coverage and funding to 880,000 children by 2001. Child Care and Development Block Grants to states provided federal funding to subsidized childcare costs for low-income families, with approximately 15% of eligible families receiving subsidies.	Almost all child care centers were regulated or licensed by their state. Regulations involved health and safety standards, staff-child ratios, maximum number of children per group, and nutrition. Centers had at the least one inspections per year.
Source	World Bank, International Comparison Program database	World Bank Poverty and Equity database	World Bank Education Statistics	OECD, Starting Strong II	OECD Family Database	World Bank Education Statistics	OECD reports	OECD reports

Note. (a) Based on purchasing power parity converted to constant 2011 international dollars, therefore suitable for cross-country comparison (b) The gross enrollment rate is the ratio between all students enrolled in pre-primary education, regardless of age, and the population of official pre-primary education age. As over- and under-aged students are included this ratio can exceed 100%. (c) The net enrollment rate is the ratio between all students in the theoretical age group for pre-primary education enrolled in that level and the total population in that age group. (d) Rates do not reveal the informal or unlicensed arrangements. (e) For Canada, the enrollment rate for children aged 0-5 years is 24%. Data on the enrollment rate for children aged 0-3 years are not available. (f) obtained by dividing the total number of enrolled pupils at pre-primary education (i.e., typically children at least 3 years old and not older than 6) by the number of full-time equivalent teachers at that level of education and in similar types of institutions. When feasible, the number of part-time teachers is converted to 'full-time equivalent' teachers.

Second, answering calls to increase replication efforts in the field, we used the same statistical procedures across the seven studies, and exploited very similar measurements of our key variables, allowing us to pinpoint if and where differences in the association between time in center care and behavior problems were due to actual differences in such relation and not due to disparities in study design or data analysis (Duncan et al., 2014). We used meta-analytic approaches to consider both the variations and average effects across studies.

Third, to address internal validity concerns, this study examined not only between-child associations, but also within-child variation employing individual fixed effects. A fixed effects approach allowed us to rule out between-child differences as a source of selection bias. Moreover, this approach directly tested the hypothesis that increases in hours in center-based care predict increases in externalizing behavior problems for toddlers and preschoolers.

For the seven studies, we examined both linear and nonlinear associations between hours in center care and externalizing problems, with the aim of testing whether the association might potentially become increasingly large at an increasingly higher number of hours in center care. We also examined family income and maternal education as possible moderators of associations between hours in center care and externalizing problems. These moderators were examined given evidence that children from more or less socioeconomically advantaged families may be at greater or lesser risk given time in center care. For example, the behavioral risks of center-based care may be particularly strong among middle class families in the United States (Huston et al., 2015), a finding that is similar to that observed in Canada where scaling up universal center-based care led to increases in problem behavior for most children but decreases for those from disadvantaged households (Kottelenberg & Lehrer, 2017; also see, Côté et al., 2007). On the other hand, scaling up center-based care in Germany predicted the greatest increases in problem behavior for children from disadvantaged versus advantaged households (Felfe & Zierow, 2018). Taken together, these

findings – although mixed with regard to the direction of moderation – indicate that family socioeconomic factors may alter the risks of center-based care for externalizing problems.

Method

We analyzed seven longitudinal datasets from five different countries. We selected these studies based on i) access to the data, ii) studies comprised of at least two measures of center-based care quantity, as well as at least two repeated measures of externalizing behavior problems, either broadly defined or restricted to measures of physical aggression (measured contemporaneous with quantity), and iii) the aim to include studies from varying sociopolitical contexts. Table 2 provides descriptions of these datasets.

Table 2

Description of the Seven Longitudinal Studies

	EMIGARDE	QLSCD	BIKS	Pre-COOL	BONDS	FLP	NICHD SECCYD
Country	Canada	Canada	Germany	Netherlands	Norway	United States	United States
Sample (total N)	515	2,223	554	~3,000	1,157	1,292	1,364
Number of assessments used / Time points (in children's age in months)	24, 36, 48	18, 30, 42, 48	43, 69	28, 42	24,36,48	35, 48, 58	24, 36, 54
Year of birth	2003 -2004	1997-1998	2002-2003	2008	2006-2008	2002-2004	1991
Population represented by the sample	Children born mothers who had taken part of a pre-natal study in four Montreal maternity hospitals.	Provincial cohort of children born to mothers residing in Quebec.	Children born in eight counties in two federal states (Bavaria and Hesse).	Children enrolled in daycare/preschool centers in specific areas (center-based sample).	Children born in five municipalities in southeast Norway.	Random sample of children born in six low-income rural counties in SE North Carolina and in central Pennsylvania.	Children who were in center-based care the year before pre-K at the 10 locations data were collected.
Outcomes							
Externalizing problem behaviors	SBQ - Aggressiveness, (mother rating; α =.73, .76, .65)	SBQ - Aggressiveness, (mother rating; α =.78, .81, .82, .77)	SDQ - Aggressiveness (teacher rating; α = .88, .90)	BITSEA – behavioral problems scale (teacher rating; α = .84, .85)	Self-made Aggressiveness (teacher rating; α = .85, .87, .89)	SDQ Aggressiveness (teacher ratings α = .75, .68; maternal rating α = .63, 68)	CBCL– externalizing behaviors (teacher rating; α =.90, .91, .92; mother rating; a = =.90, .91, .92)

Note. EMIGARDE=Life experiences and psychosocial development of the child: the role and quality of child care services study; QLSCD=Quebec Longitudinal Study of Child Development; BIKS-3-10 = Educational Processes, Competence Development and Selection Decisions in Preschool and School Age; Pre-COOL = Cohort Research on Educational Careers – young child; BONDS = Behavior Outlook Norwegian Developmental Study; FLP = The Family Life Project; NICHD SECCYD = National Institute of Child Health and Human Development Study of Early Child Care and Youth Development.

Studies and Samples

Life Experiences and Psychosocial Development of the Child: The Role and Quality of Child Care Services Study (EMIGARDE), Canada

Participants for the EMIGARDE study were drawn from a prenatal–perinatal study of births between June 2003 and April 2004 conducted in four Montreal hospitals. Participants were contacted and invited to participate in a follow-up study concerning the development of their child prior to the children's second birthday (for details, see Côté et al., 2013). This study provides detailed information about family (socio-demographics), child, and the child care characteristics, measured through parental interviews (mostly with mothers) when children were 24, 36, and 48 months old.

Longitudinal Study of Child Development in Quebec (QLSCD), Canada

The QLSCD, also known as ÉLDEQ for its acronym in French, started in 1998 with a sample of 2,120 children representative of a provincial cohort of children in Quebec (visit https://www.jesuisjeserai.stat.gouv.qc.ca/ for more information). The first phase followed children annually starting when they were 5 months old until they were about 4 years with the goal of gaining a better understanding of the factors related to development of basic skills needed for educational success. A wide range of family, parenting, and child care data were collected through interviews with families.

Educational Processes, Competence Development, and Selection Decisions in Preschool and School Age Study, BiKS-3-10, Germany

The BiKS-3-10 study started in 2005 with an initial sample of 547 children who were attending 97 different preschools (for details see Weinert, 2013). Starting in the fall of 2005, data were collected at intervals of about half a year completing six times of measurement (before school entry), including assessments and parental- and teacher interviews. The study

collected data on children's externalizing behavior during wave 3, when children were 43 months old and wave 5 when children were 69 months old.

Pre-COOL, The Netherlands

Pre-COOL is a national cohort study including in total over 3,000 children born in 2008. Pre-COOL includes a "center-based" sample, recruited through daycare centers and preschools close to COOL schools (see Veen et al., 2012; for further details). The sample was drawn by Statistics Netherlands (CBS). While the sample is not fully representative of all Dutch children, the sample is well-spread geographically and across urban, semi-urban and rural areas. Children were assessed from age 2 until the end of primary school. For this study, only the first two waves (age 2 and 3) were used as essentially all children in the Netherlands start kindergarten, which is considered part of the school system, when they turn 4.

Behavior Outlook Norwegian Developmental Study (BONDS), Norway

The BONDS is an ongoing longitudinal study, which includes 1,159 families recruited from (almost universally attended) child health clinics in five municipalities in southeast Norway, when the children were 6 months old (participation rate was 60%). Two families later withdrew their participation. Children were born in 2006, 2007, and 2008. Interviewer-based data collections (at age 6 and 12 months, thereafter annually through school entry) included a wide range of measures of child behaviors and skills, observed parent-child interactions, and parenting. Questionnaires to center care teachers, including behavioral measures, were administered at 24, 36, and 48 months). For further information, see Nærde et al., 2014.

Study of Early Child care and Youth Development (SECCYD), United States

The SECCYD is a multi-site longitudinal study of including 1,364 children born in 1991 (see http://secc.rti.org/ for more information). Families were recruited from 10 different sites across the United States, and although it was not nationally representative, the study

population was diverse and included children from different socioeconomic backgrounds and was novel in its extensive assessments of child care use. The wide range of data on child care use and family data were collected through interviews with and observations of families and teachers, and observations done in 2-hours visits to homes and child care settings when children were 24, 36, and 54 months old.

The Family Life Project (FLP), United States

The FLP was designed to study young children and their families in two of the four major geographical areas of the United States with high poverty rates (Dill, 1999)—southeast North Carolina (NC) and central Pennsylvania (PA; for details see https://flp.fpg.unc.edu/). Specifically, three counties were sampled in each region to reflect African-American families in the South and Caucasian families in Appalachia, respectively. Sampling was based on a developmental epidemiological design in order to recruit a representative sample of 1,292 children whose families resided in one of the six counties at the time of the child's birth. Low-income families in both states and African-American families in NC were over-sampled. African-American families were not over-sampled in PA given the small representation of African-American families in this region (< 5%; Vernon-Feagans and colleagues, 2013). A wide range of family, parenting, child care, and community data were collected through interviews, observations, and direct assessments conducted in the child's home and child care settings when children were 35, 48, and 58 months old.

Measures

Externalizing behavior

We used teacher externalizing problems reports in all datasets that had this information available, given evidence that quantity-effects are most consistently detected when using these ratings (NICHD ECCRN, 2002). However, teacher reports were not available in all the studies, thus, when not available we used parent reports. The EMIGARDE

and the QLSCD studies used the Social Behavior Questionnaire (SBQ; Tremblay, 1991) to assess children's behavioral problems. The scale was comprised of 6 items, that included behaviors such as "hits, bites, kicks other children" and "fights more". The BIKS and FLP studies used the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), a brief teacher-report questionnaire in which the child is rated on various domains of behavior. The SDQ-externalizing problem behaviors items include behaviors such as "fights or bullies other children" and "can be spiteful to others". In pre-COOL, externalizing behavior was measured using five items (e.g., "restless and can't sit still," and "hits, bites, or kicks") from the behavioral problems subscale of the Brief Infant Toddler Social and Emotional Assessment (BITSEA; Briggs-Gowan & Carter, 2002). The BONDS study used a frequency-rating of eight aggressive behaviors, including "hits other children" and "pulls hair", ranging from 1 (never, not in the past year) to 7 (3 times daily or more). The measure was completed by both parents and child care teachers (see further details in Dearing et al., 2015 and in Nærde et al., 2014). The SECCYD study used the Child Behavior Checklist (CBCL; Achenbach, 1991) completed by parents and child care providers. The parent CBCL includes 113 problem behavior questions and the teacher CBCL lists 100 questions addressing internalizing and externalizing problems. The Externalizing Problems Scale includes items such "child argues a lot" that are rated as not true (0), somewhat true (1), or very true (2) of her child. Scores of the scale items are summed and converted into standard T scores, based on normative data for children of the same age.

Child care quantity

In most studies, care quantity was measured by the number of hours per week that children attended center-based care excluding any other types of care arrangements (i.e., home care by a parent, group care by a relative or a nonrelative) reported by the main caregiver (most of the times mothers) at each time point. In the particular case of the pre-

COOL, the study collected the number of days that children attended center-based care (including half days), which we transformed to weekly hours. To avoid estimates biased by outliers, we truncated the quantity of care variable on the 95th percentile. Across datasets, we set the quantity of time in center-based care to zero for children attending other types of care arrangements at any particular time point. We also parameterized the units as 10th hour of care to make the coefficient easily interpretable. For samples where the quantity of care was also reported for other types of child care arrangement, we presented secondary analyses in the supplementary materials (Table S1).

Covariates

Because our main goal was to test the association between within-person changes in center care quantity and externalizing behavior, we included a restricted set of covariates to account for selection, and primarily relied on time-varying covariates for this purpose. These covariates included single parenthood, total number of siblings or new siblings in the family, and parental employment, measured at all time-points. The selection of covariates was pragmatic; we included those available in all datasets. Additional time-invariant covariates such as child's gender (49% female), race/ethnicity, and immigrant status were included in the between-person models. We were only able to include race/ethnicity in the two samples from the U.S. (57% and 80% White; 42% and 13% African-American; 1.2% and 5% Latinx). All other study samples were from countries that do not allow these demographic indicators to be collected. Table 3 presents descriptive statistics of outcome and predictor variables in the analytic samples, Table S2 in the supplementary materials presents descriptive statistics of the covariates in each sample.

Table 3

Descriptive Statistics for Outcome and Predictor Variables in the Analytic Samples

	EMIGARDE		QLSCD		BIKS		Pre-COOL		BONDS		FLP		NICHD SECCYD	
Variables	*M* (*SD*)	min, max	*M* (*SD*)	min, max	*M* (*SD*)	min, max	*M* (*SD*)	min, max	*M* (*SD*)	min, max	*M* (*SD*)	min, max	*M* (*SD*)	min, max
Outcome: ***Aggressive behavior***														
Externalizing T1	1.88 (1.84)	0, 9.17	1.33 (1.26)	0, 7.9	0.9 (0.53)	0, 2.7	2.09 (0.79)	0, 5	1.78 (0.78)	1, 4.5	0.63 (0.42)	0, 2	47.38 (9.94)	30, 82
Externalizing T2	1.53 (1.68)	0, 9.17	1.75 (1.4)	0, 10	0.74 (0.49)	0, 2.7	1.99 (0.72)	0, 4.4	1.9 (0.89)	1, 5.83	0.44 (0.34)	0, 2	45.89 (10.03)	30, 80
Externalizing T3	1.94 (2.12)	0, 10	2.23 (1.54)	0, 9	--	--	--	--	1.6 (0.76)	1, 6	0.39 (0.36)	0, 1.8	50.21 (9.56)	35, 88
Externalizing T4	--	--	2.04 (1.46)	0, 9.5	--	--	--	--	--	--	--	--	--	--
Main Predictor: ***Quantity of center-based***														
Hours in center care T1	27.16 (17.62)	0, 50	6.17 (13.34)	0, 40	24.61 (7.36)	0, 38.6	13.15 (8.75)	0, 50	29.11 (13)	0, 42	11.45 (21.70)	0, 94	6.2 (14.03)	0, 45
Hours in center care T2	27.73 (16.45)	0, 45	9.78 (16.46)	0, 45	27.93 (7.12)	0, 40.7	12.15 (7.74)	0, 50	32.21 (10.27)	0, 42	15.03 (19.67)	0, 97	9.57 (15.86)	0, 45
Hours in center care T3	28.61 (17.22)	0, 50	15.78 (18.35)	0, 45	--	--	--	--	33.25 (8.64)	0, 42	16.07 (19.32)	0, 98	16.7 (15.65)	0, 45
Hours in center care T4	--	--	17.48 (18.46)	0, 45	--	--	--	--	--	--	--	--	--	--
Hours in center care (*between SD)*	(13.86)	--	(13.150)	--	(6.31)	--	(8.39)	--	(7.43)	--	(11.56)	--	(12.38)	--
Hours in center care (*within SD)*	(11.43)	--	(10.37)	--	(3.81)	--	(2.76)	--	(3.93)	--	(15.41)	--	(10.58)	--

Moderating variables

To test the interaction with family background, we included maternal education (a dichotomous measure indicating whether the mother has secondary education or more) and family income, as available in the individual datasets.

Missing Data

Missing data was mostly due to item nonresponse and to attrition, with missingness ranging from 7 to 27% in the outcome variables and from 0 to 32% in covariates. Following best practice for handling moderate to large amounts of missing data, we used multiple imputation with chained equations when using Stata (ICE in Stata, Royston & White, 2011) and Proc MI when using SAS to generate 20 datasets, using all variables described in Table S2 as covariates in the imputation model.

Analytical Approach

In this confirmatory study, we estimated random effects models that account for the between-person variation of intercepts and slopes around the population average trajectory. Although the random effects models allowed us to include time-invariant variables (i.e., sex, ethnicity), their identifying assumptions of conditional independence, i.e., the unobserved heterogeneity (or random effects) is uncorrelated with other observed variables and is difficult to meet. Thus, to mitigate the risk of omitted variable bias, we then employed a more conservative approach. Our second analytical strategy was within-person fixed effects analysis. Fixed effects analyses provide estimates of the within-child changes in externalizing behavior as a function of within-child changes in center care quantity (conditioning on changes in covariates).

In considering the two modeling approaches, it is also helpful to note that the random effects estimates capture all between- and within-child variation, answering the question: are higher *levels* of time in center care associated with higher *levels* of externalizing problems?

Thus, the random effects estimates allow comparisons among all children, including those that never enter center care. However, as mentioned, unobserved between-child heterogeneity is a serious concern for the random effects estimates, but remedied by the fixed effects estimates. Compared with the random effects focus on levels of time in care, the fixed effects estimates answer the question: are within-child *increases or decreases* in hours in care associated with within-child *increases or decreases* in externalizing problems?

The resulting fixed effects equation can be written as follows:

$$y_{it} - \bar{y}_i = \beta_x(x_{it} - \bar{x}_i) \quad (1)$$

In our models (ignoring covariates and error terms), y_{it} is the population-averaged (PA) score for child *i* at time *t,* and $\bar{y}_i$ the average PA score across time points. Likewise, x_{it} and $\bar{x}_i$ are center care quantity for child *i* at time *t* averaged across all time points. As such, β_x is interpreted as the average within-child association between center care quantity and individual child externalizing behavior. The fixed effects model was then expanded to include time-varying covariates to account for time-varying confounders. We included the same time-varying covariates in the seven studies to guarantee we were comparing the same models across all studies. Although fixed effects estimates, by design, control for all possible time-invariant sources of bias, unmeasured time-varying factors may still bias estimates, and a correctly specified model is contingent on the inclusion of all probable time-varying confounders.

For all models, we also estimated nonlinear associations by adding a quadratic term for center care hours in the equation. In addition, as a third step, we tested potential interactions with family background, driven by concerns that negative effects of the hours in center-based care were primarily found among non-Hispanic, White, middle class children (Huston et al., 2015). We included maternal education and family income as moderators

allowing the within-child estimates of time in center care to vary in magnitude along these socioeconomic dimensions. Each potential moderator was examined in separate analyses.

In the final step of our primary analysis, we estimated the overall effect across the seven studies using meta-analysis, computing the fixed effects weighted (by sample size) average of the studies' individual coefficients. Fixed effects meta-analysis assumes that all studies are estimating a common effect, in our case the association between hours spent in center-based care and externalizing behavior. This assumption implies that observed variation among the different study estimates is due to measurement error or differences in sampling procedures, not to "real" differences (Riley et al., 2011).

Sensitivity and generalizability analyses

To examine the robustness of our results, we estimated several alternative models to those presented in the manuscript. Presented in the supplementary materials, we estimated the following sensitivity checks: (1) in five of the seven datasets that collected information on child care arrangements other than center-based care, we examined any form of nonparental care (rather than center-based care, exclusively) as the "treatment" of interest, with exclusive parental care serving as the counterfactual; (2) we calculated a cumulative index of hours in care across the available time points measured (i.e., a child with constant 30 hours of care across three time points would have a cumulative load of 30, 60, and 90 hours in our models), and used this cumulative index in fixed effects analyses (please see the supplemental materials for further details on the cumulative hours models); (3) to further examine potential non-linearities, we estimated within-child and between-child differences in no center care (0-4 hours), part-time center care (five to 25 hours of care), and full-time center care (more than 25 hours of care); (4) we reran the analyses without maternal employment as a covariate, given the possibility it could capture some of the center care effect; and (5) in four of the seven datasets with available time-varying measures of family income, we included this

variable as a time-varying covariate and potential moderator. In addition to these alternative modeling approaches, we also include (in our supplemental materials) analyses that inform the generalizability of our results; given that neither our random effects nor our fixed effects models can inform what would have happened if children who never entered center care did, in fact, enter this form of care (and, similarly, for any child who was in center care but was always in a stable amount of care, what would have happened if amount of care had changed), we examined differences in background characteristics between these groups of children.

Results

In Table 4, for both random and fixed effects models, we first present linear associations, followed by nonlinear estimates and moderated effects. Overall, we found little evidence that increases in time spent in center-based care predicted increases in externalizing problems in any of the random effects or fixed effects models across the seven studies.

For the linear estimates, evidence of statistically significant or approaching significant results were evident in two of the seven studies, but only in the random effects models. For the Pre-COOL study in the Netherlands (b = .140, SE = .034, p <.001) and the BONDS study in Norway (b = .036, SE = .020, p <.10), more hours in center care predicted more externalizing problems, albeit only statistically significant at 10%-level for the Norwegian study. However, in the fixed effects models, none of the datasets revealed statistically significant associations between time in center-based care and externalizing problems, and nearly all of these associations were effectively zero.

Table 4

Standard Deviation Changes in Externalizing as A Function of 10hr Increase in Time in Center-Based Care

	EMI-GARDE parent-report	**QLSCD** parent-report	**BIKS** teacher-report	**Pre-COOL** teacher-report	**BONDS** teacher-report	**FLP** teacher-report	**NICHD SECCYD** teacher-report
	b (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)	*b* (*SE*)
Random effects							
Quantity (linear estimate)	.003 (.020)	.007 (.008)	.093 (.058)	.140** (.034)	.036^{+} (.020)	.006 (.011)	-.041 (.031)
Nonlinear specification							
Quantity	.009 (.023)	-.004 (.010)	.098 (.059)	.146*** (.032)	.070* (.033)	0.000 (.014)	-.084 (.060)
Quantity- squared	.006 (.015)	.010+ (.006)	.024 (.041)	.024 (.019)	.019 (.016)	0.003 (.005)	.029 (.022)
Interaction with maternal education							
Quantity	.008 (.064)	-.002 (.015)	.137^{+} (.071)	.118* (.057)	.058^{+} (.031)	.010 (1.057)	-.052 (.045)
Education	-.440** (.156)	.030 (.058)	-.115 (.083)	-.154^{+} (.085)	.005 (.107)	-.297*** (.060)	-.231*** (.061)
Education x Quantity	-.006 (.067)	.012 (.016)	-.107 (.104)	.027 (.055)	-.045 (.044)	.009 (.024)	.016 (.030)
Interaction with family income							
Quantity	-.063 (.061)	.005 (.008)	.093 (.057)	.106 (.070)	.032 (.020)	.002 (.009)	-.046 (.050)
Income	-.019 (.020)	.098* (.046)	.001 (.009)	-.023 (.013)	.072 (.072)	-.099*** (.015)	.004 (.011)
Income x Quantity	.009 (.009)	.008 (.017)	-.001 (.009)	.005 (.008)	.038 (.069)	.002 (.006)	.001 (.006)
Fixed effects							
Quantity	.032 (.059)	-.016 (.027)	.065 (.109)	.069 (.110)	.010 (.046)	.001 (.011)	-.051 (.053)
Nonlinear specification							
Quantity	.110 (.099)	.044 (.073)	.065 (.110)	.071 (.114)	-.004 (.050)	-.002 (.014)	-.086 (.087)
Quantity- squared	-.042 (.042)	-.017 (.021)	.005 (.080)	-.003 (.045)	-.017 (.028)	.002 (.006)	.012 (0.29)
Interaction with maternal education							
Quantity	-.121 (.184)	.002 (.064)	.132 (.129)	.069 (.157)	.021 (.071)	.001 (.023)	-.046 (.093)
Education x Quantity	.169 (.195)	-.022 (.068)	- .180 (.190)	.000 (.135)	-.019 (.088)	.007 (.024)	-.007 (.094)
Interaction with family income							
Quantity	-.127 (.209)	-.024 (.028)	.076 (.106)	.035 (.165)	.022 (.049)	.002 (.011)	-.089 (.075)

Income x Quantity	.022 (.028)	.050 (.074)	.009 (.017)	.005 (.022)	-.091 (.132)	.002 (.006)	.009 (.013)

Note. All multiple-imputation models are based on 20 imputed data sets. Coefficients for hours in care are reported with 10-hour increments. Fixed effects models controlled for time-varying number of siblings, single parent, maternal employment status. Random effects models controlled for time-varying maternal employment status, number of siblings, and single parent. Additionally, we included maternal education (at time point 1), parents' migrant status (or ethnicity in the case of the U.S.), sex of the child, and family income. $^{+}$ $p<.10$, * $p<.05$, ** $p<.01$, *** $p<.001$.

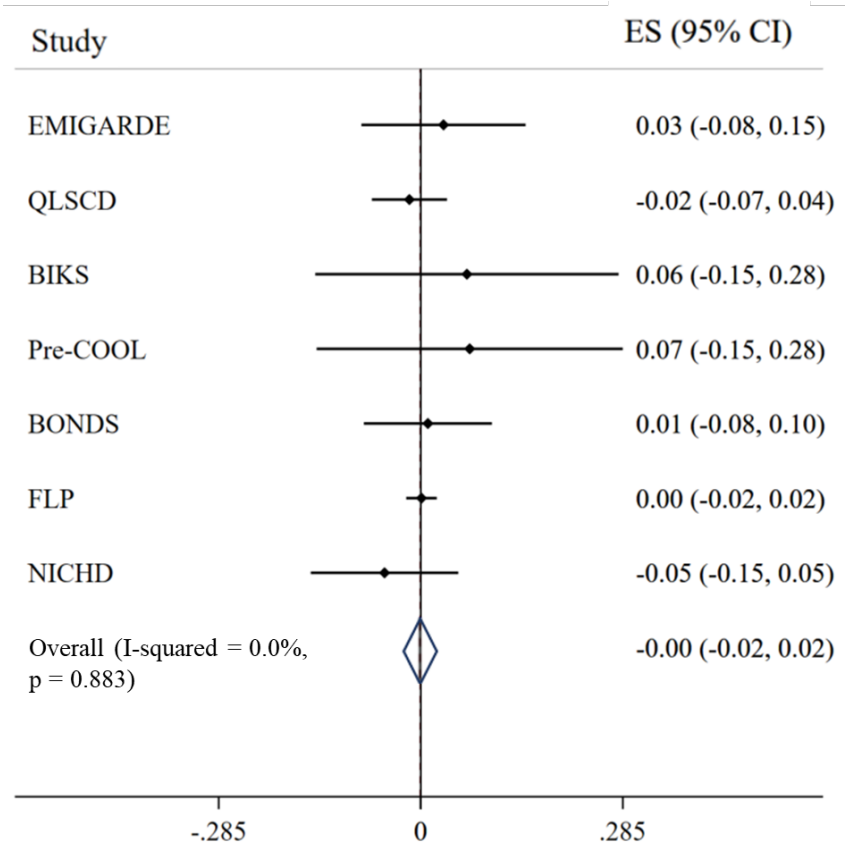
With regard to potential non-linear effects of hours in center care, we were interested in testing whether the association with externalizing problems could be increasingly large at increasingly higher levels of hours in center care. To examine this possibility, we estimated a quadratic specification for hours. Yet, as can be seen in Table 4, the quadratic specification did not reach statistical significance ($p<.05$) in any of the seven studies for either the random effects or the fixed effects models.

In regard to the moderators we tested, there was no evidence of family background moderation of quantity effects. The magnitude of the effect sizes did not vary along the two SES dimensions examined, i.e., family income and maternal education. This was the case both in the random and fixed effects models.

For the linear fixed effects estimates, we conducted a meta-analysis of results from the seven datasets. Figure 1 displays: (1) effect sizes and confidence intervals for each study and (2) the meta-analytic effect size for fixed effects estimates of the association between changes in hours of center care and changes in externalizing problems. In the figure, small black diamonds indicate individual study effect sizes, horizontal lines depict confidence intervals, and the larger (unfilled) diamond represents the meta-analytic effect estimate. Not surprisingly given the very small individual effect sizes, the meta-analytic association between hours spent in child care and externalizing problems was .00, and null.

Figure 1

Forest Plot of the Inverse-Variance Weighted Fixed-Effect Meta-Analysis of the Relation of Hours in Care and Behavior Problems Across the Seven Studies



Sensitivity and Generalizability

To examine the robustness of our results, we estimated several alternative model specifications, including: (1) time in nonparental care as our "treatment" (and time in parental care as the counterfactual; Table S1), (2) the effects of accumulating hours in center care (Table S3), (3) non-linear effects of hours in center care using categorical indicators (Table S4), (4) center care effects without controlling for maternal employment (Table S5), and (5) time-varying income (Table S5). In sum, we find consistent evidence that our primary model results were robust to these alternative approaches to detecting associations between time in center care, or any form of nonparental child care.

Across all of these models, the only alternative specification that appeared to produce results that differed from our original models was when using a categorical approach to testing non-linearities (see Table S4). Yet, even here, the results were mixed, at best, and contradictory. In the random effects model for one of the seven datasets, children never in center care had fewer problems than those in full-time center care, and in two of the seven datasets children in part-time center care had fewer problems than those in full-time center care. However, these differences were not evident in any of the datasets when estimated in the fixed effects models. In fact, in the fixed-effects model for one of the seven datasets, moving from no center care to full-time center care predicted *decreases* in problems, although significant only at a 10%-level.

Finally, to more carefully consider to whom our results might generalize, we compare (Table S6) covariate balance for children with stable hours in care (i.e., never entering care or remaining in stable number of hours of care) with children who were in varying hours of care. We did so for center care and any form of nonparental care, center or otherwise. The number of children who were stable varied across datasets from about 10% to more than 33%, but in all cases these percentages were dominated by children who were never in center care (or never in nonparental care); fewer than 5% of children in all datasets entered center or nonparental care and then maintained a stable number of hours. Not surprisingly, given links between maternal employment and child care, children who remained stable differed from other children in most datasets in one or more of the following areas: mothers being less likely to be employed, families having lower incomes, and families having a greater likelihood of single parenthood. In some datasets (and depending on whether center care or any nonparental care was the focus), maternal education and family immigration status also emerged as significant differences between these groups of children. We discuss these results as they pertain to the generalizability of our primary models of interest.

Discussion

Decades of debate on whether extensive time spent in center-based care might pose a risk for developing externalizing behavior problems has yielded inconclusive and conflicting findings. In part, variations in statistical methodologies across studies have complicated the ability to form firm conclusions, with concerns that child and family selection effects may be a frequent source of bias in this non-experimental field of study. And, on the other hand, a lack of variation in sociopolitical contexts has made it difficult to generalize widely, because the field has relied very heavily on U.S. studies. With limitations to the current state of the field in mind, we extended the evidence base by examining the association between hours spent in center-based care and behavior problems using data from seven longitudinal studies conducted in five countries. In doing so, we examined both between child and within child variation in center-based care use.

Across the seven studies, our results were unequivocal: we found little evidence that time spent in center care relates to more externalizing behavior problems. Only one of the seven estimates from the random effects model reached statistical significance at conventional 5%-level –from the Pre-COOL study in the Netherlands–and none of the fixed effects estimates examining the within-child association between changes in hours and changes in behavior problems were significantly different from zero. Moreover, we found no evidence that indicators of SES (i.e., maternal education, household income) moderated the effect of time in center care.

Contributions of the Current Study to the Existing Literature

Two contributions of the current study are particularly noteworthy, the first relating to internal validity and the second relating to external validity. With regard to internal validity, our results support arguments (e.g., Dearing et al., 2015; McCartney, et al., 2010; Zachrisson et al., 2013) that associations between quantity of nonparental care, center-based or

otherwise, and externalizing problems is not robust to conservative estimation approaches that account for unobserved selection factors. With regard to external validity, we find consistent (null) effects across seven studies, spanning five countries on two continents, some of them targeted at socioeconomically disadvantaged families (BIKS, Family Life), and others with a skew towards (but by no means restricted to) more affluent families (BONDS, SECCYD). Here, we elaborate on these points.

Internal validity

In our study, we address internal validity by relying on within-person fixed effects analyses to account for unobserved selection bias. Relatively few previous studies have taken this particular approach. Yet, those that have, have found results consistent with ours, including two studies using the NICHD SECCYD (McCartney et al., 2010; Morrissey, 2009), one of the seven datasets in our study. The current study extends these findings not only by including a second longitudinal U.S. sample, but also five other samples from Canada and Europe. Also consistent with our findings, Zachrisson et al. (2013) reported null effects of nonparental child care hours on externalizing problems using a different Norwegian sample than we have examined, and using maternal reports of externalizing problems. Thus, collectively in the literature, we now found no evidence of a within-child relation between hours in care and behavior problems in two U.S. samples, two Norwegian samples, and samples from Canada, Germany, and the Netherlands.

Other studies using various quasi-experimental designs (e.g., difference-in-differences, instrumental variables, and sibling fixed effects) to account for unobserved selection have also primarily produced null results (for review, see Dearing & Zachrisson, 2017). The only notable exception is the German study of increases in full-time slots (Felfe & Zierow, 2018), preventing us from drawing unambiguous conclusions about the absence of strong identification strategies being the (only) reason for positive associations between

quantity of care and externalizing problems in previous studies. Moreover, a set of studies from Canada comparing externalizing problems for children in Quebec with those in other provinces, before and after an ECEC reform that led to more children in Quebec attending child care; these studies report that the reform led to higher levels of problems among children in Quebec (e.g., Baker et al., 2008). While questions have been raised about the quality of ECEC in Quebec during these reforms, it is also important to note that the counterfactual condition is not the same child at a different time with more or less exposure to ECEC, but a comparison of children from different birth cohorts with different likelihoods of attending ECEC. While informative, the Quebec results do not address the question of child care quantity per se. Regardless, the Quebec results appear to the be one of a few exceptions in a literature of null results from quasi-experimental tests of the hypothesis that exposure to ECEC causes problem behaviors. With a similar counterfactual (of not being exposed to center-based childcare), Berger, Panico, and Solaz (2021) found early entry in France to increase levels of behavior problems at age 2, a finding similar to that reported by Dearing et al. (2015), who had a longer follow-up and found the initial spike in problem behavior following early entry into center care to decrease to zero at age 4.

External validity

A second important contribution of the current study is in the realm of external validity. To date, the literature on the topic had been dominated by studies in the United States with a few exceptions from Canada, European countries, Australia, and Japan. When we consider exclusively studies with relatively strong internal validity, the research base is even more limited, to the U.S., Norway, and Germany (also including Canada and France if we consider studies on ECEC and externalizing problems more broadly). The heavy reliance on U.S. evidence has called into question the generalizability of findings to other sociopolitical contexts with different ECEC and family policies. With the purpose of

addressing external validity issues and recognizing the key role of replication for increasing the confidence of robustness of findings (Duncan et al., 2014), we brought together seven longitudinal studies to replicate former studies using a similar statistical approach across them. As is evident from Table 1 in the introduction, all of these datasets come from high-income countries. Yet, the table also shows that the countries cover a wide range of sociopolitical contexts, from the most (Norway) to one of the least (U.S.) progressive child- and family-policies among high-income countries (Chzhen, et al., 2019). While scholars have speculated that the potentially negative effects of center-based child care may be less pronounced in contexts with more comprehensive welfare and health-policies (Dearing & Zachrisson, 2017), this does not receive support from our results; they are consistent across sociopolitical contexts.

Moreover, apart from the sociopolitical contexts, our results are strikingly similar despite considerable differences in designs and samples across studies. These include the mean and variability in hours of care, the number of assessments, different externalizing measures including some reported by the mother and some by teachers, and—as mentioned—samples from different strata of the populations. Additionally, the null associations between time spent in center care and externalizing behaviors were also consistent regardless of children's age or the time between observations in the different samples. This further strengthens the argument that time in child care is not detrimental for children in different developmental periods nor while considering other contextual and measurement issues.

In addition, our results replicate when all types of nonparental care (not exclusively center-based care), non-linear associations, cumulative time, and transitions from zero- to part-time to full-time are considered, as shown in the Supplementary materials. Despite all of these between-sample differences our findings were conclusive in that time spent in center-based care does not pose a risk for developing externalizing behavior problems. The slight

variations in studies that make replications closely similar but not identical broadens the generalizability of the results.

Still, the question of whether there might be differentiated quantity effects for particular subgroups in the population remains. Some studies have found that this could be the case primarily for White non-Hispanic middle-class children who exhibited behavior problems more consistently than African American and Hispanic children after attending child care (see Huston et al., 2015 for review), although this conclusion was heavily driven by results from the NICHD SECCYD. Moreover, this evidence comes mainly from studies in the United States. We found no evidence that the association between center-based care and externalizing problems differed as a function of family income or parent education, despite considerable socioeconomic variation in each of the seven samples. This study extends generalizability of the null finding across subgroups given that our samples include a considerable range of children across the socioeconomic spectrum. In sum, our results were consistent in that we did not find evidence for quantity effects in these particular populations of White middle-class children living in high-income countries.

Study Limitations

In interpreting our findings, it is important to note that we examined only short-term effects and are unable to address whether longer term harm might emerge. This has, for instance been found in a Canadian study using one of the datasets included in the current study (QLSDC), yet notably with a research design which does not account for unobserved selection (Pingault et al., 2015), as well as in a quasi-experimental Canadian study (Baker et al., 2019). The opposite pattern was observed in a Norwegian study looking at age of entry into center care, which also applied one of the datasets included in the current study (the BONDS); short term increases in behavior problems rapidly declined with age (Dearing et

al., 2015). In addition, the consistency in our results despite some variations in the externalizing problems measures underscores the robustness of our null findings.

However, we must also note that while our fixed effects models have important strengths, we cannot rule out unobserved time-varying potential confounds; changes in hours may have been driven by unobserved time-varying family selection factors. We attempted to account for potential time-varying sources of bias by including controls for single parenthood, new siblings in the family, and parental employment, measured at all time-points. Given the fact that the results were unequivocally null across studies, we have greater confidence that at least in this sample of studies there is not an association between time in center care and externalizing problems. Yet, time-varying omitted variables may bias results upward or downward, and remain a reasonable concern for our results. Moreover, the fixed effects approach we use in this study does not account for the possibility that there are cumulative effects of exposure to high quantities of child care. If exposure in one time period has effects on behavior problems in a time later period, beyond concurrent changes in child care quantity, we would not capture this effect with our design.

Another limitation of the current study is that none of these samples were nationally-representative, although, they represented different populations across the SES distribution. Moreover, our samples vary not only in sociopolitical contexts but also in historical contexts confounding these two contexts. We also must recognize that sociopolitical variations across these countries could affect the very meaning of variables such as education and income – our potential moderators of interest – a caveat worth considering for any general statements about the null findings for these moderators. Yet, on the other hand, it is worth calling attention to the fact that all of these samples were from higher-income countries; it is questionable as to how diverse these countries truly are in terms of ECEC experiences of children in care and the moderators we tested. Further research needs to explore whether

these results might generalize to children living in sociopolitical contexts that are different from those in high-income countries.

Another important consideration for interpreting our results is to compare the random effects and fixed effects estimates. The random effects estimates captured all between- and within-child variation, ultimately answering the question: were higher *levels* of time in center care associated with higher *levels* of externalizing problems? The fixed effects estimates answered the question: are within-child *increases or decreases* in hours associated with within-child *increases or decreases* in externalizing problems? The answer to both questions was overwhelming no. The fixed effects estimates are our preferred results, given the ability to rule out unobserved between-child heterogeneity, but we must note that neither estimator allowed us to examine what would have happened if children who never entered center care in these studies had, in fact, entered care, because the random effect estimates are based on levels and children who were never in center care drop out of the fixed effects analyses. Because children from disadvantaged backgrounds (i.e., unemployed parents, low income, and single-parent households) were over-represented among those who never entered center care, an added measure of caution is needed when generalizing our results to these children.

Finally, we acknowledge that absence of evidence is not the same as evidence of absence. The logic of null-hypothesis testing does not imply that when associations in our study do not reach conventional levels of statistical significance means that there is evidence of the absence of such an association. Rather, lack of statistical significance (e.g. $p < .05$) is not support for the null hypothesis, it is failure to reject it. It is therefore worthwhile also considering the practical implication of the effect sizes across studies. These range from -.03 to .07 SD as a function of a ten-hour change in hours in care, with the largest one also being the least precisely estimated, and most effect sizes hoovering around zero. The overall meta-

analytic effect size is also zero, supporting the conclusion that the effect sizes do not warrant concern over negative effects of ECEC quantity on externalizing problems.

Conclusion

Across seven samples taken from five countries, we found no evidence that more time spent in center-based child care poses a risk for developing behavior problems in early childhood. This is reassuring given that trends in child care use and parental participation in the labor force are likely to remain stable. And, this reassuring evidence is all the more valuable when juxtaposed alongside evidence of long-term achievement benefits of ECEC for children, some of which are experienced directly (e.g., learning stimulation in ECEC) and others more indirectly by supporting parent participation in the workforce (e.g., Bustamante et al., 2021; Morrissey, 2017). In light of this, continued research into practices and policies that ensure early child care supports the well-being of children, and families, remains a priority internationally (e.g., improving quality of care and access).

Acknowledgements

Data from the EMIGARDE and QLSCD studies are available upon request to the Quebec Institute of Statistic data center: https://statistique.quebec.ca/recherche/#/accueil. BiKS-3-10 data are publicly available at the following DOI: 10.5159/IQB_BIKS_3_10_v6. Pre-COOL data can be accessed upon request to the Dutch National Center of Expertise and Repository for Research Data (DANS; www.dans.knaw.nl). BONDS data are available upon request to the Norwegian Center for Child Behavioral Development (NUBU), Oslo, Norway: post@nubu.no or to the author Ane Nærde (mail: ane.narde@nubu.no). Data for the NICHD- SECCYD study are available through the archive at https://www.icpsr.umich.edu/web/ICPSR/series/233, however, some data might require application and IRB approval. Data for the FLP are available through the archive at https://www.icpsr.umich.edu/web/ICPSR/studies/34602/datadocumentation, however, some

data might require application and IRB approval. The analytic code necessary to reproduce the analyses is available from the first author. There are no additional materials necessary to attempt to replicate the findings presented here. The analyses presented here were not preregistered. The Family Life Project and Dr. Daniel Berry's contribution was supported by a grant from the National Institutes of Health Office: 4UH3OD023332-03; UG3OD023332-01; UH3OD023332-01, as well as the previous grants from the National Institute of Child Health and Human Development 1PO1HD39667 and 2PO1HD039667. Dr. Ane Nærde´s contribution was supported by funding from the Research Council of Norway (grant # 283438/H20). Dr. Henrik Daae Zachrisson’s contribution was supported by funding from the European Research Council Consolidator Grant ERC-CoG-2018 EQOP (grant # 818425).

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4–18 and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.

Averdijk, M., Besemer, S., Eisner, M., Bijleveld, C., & Ribeaud, D. (2011). The relationship between quantity, type, and timing of external childcare and child problem behaviour in Switzerland. *European Journal of Developmental Psychology*, *8*(6), 637–660. https://doi.org/doi:10.1080/17405629.2011.571846

Baker, M., Gruber, J., & Milligan, J. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, *116*(4), 709–745. https://doi.org/doi:10.1086/591908

Baker, M., Gruber, J., & Milligan, K. (2019). The long-run impacts of a universal child care program. *American Economic Journal. Economic Policy, 11*(3), 1–26. https://doi.org/10.1257/pol.20170603

Barnes, J., Leach, P., Malmberg, L. E., Stein, A., Sylva, K., & The, F. T. (2010). Experiences of childcare in England and socio-emotional development at 36 months. *Early Child Development and Care*, *180*(9), 1215–1229. https://doi.org/doi:10.1080/03004430902943959

Belsky, J. (1986). Infant day care: A cause for concern? *Zero to Three*, *7*(1), 1–7.

Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., & Owen, M. T. (2007). Are there long-term effects of early child care? *Child Development*, *78*(2), 681–701. https://doi.org/10.1111/j.1467-8624.2007.01021.x

Berger, L. M., Panico, L., & Solaz, A. (2021). The impact of center-based childcare attendance on early child development: Evidence from the French Elfe cohort. *Demography, 58*(2), 419–450. doi:10.1215/00703370-8977274

Berry, D., Blair, C., Ursache, A., Willoughby, M., Garrett-Peters, P., Vernon-Feagans, L.,

Bratsch-Hines, M., Mills-Koonce, W. R., & Granger, D. A. (2014). Child care and cortisol across early childhood: context matters. *Developmental Psychology*, *50*(2), 514–525. https://doi.org/10.1037/a0033379

Borge, A., Rutter, M., Côté, S., & Tremblay, R. E. (2004). Early childcare and physical aggression: differentiating social selection and social causation. *Journal of Child Psychology and Psychiatry, 45*(2), 367–376. doi.org/10.1111/j.1469-7610.2004.00227.x

Briggs-Gowan, M. J., & Carter, A. S. (2002). *Brief Infant-Toddler Social and Emotional Assessment (BITSEA) manual, version 2.0*. Yale University Press.

Bustamante, A. S., Dearing, E., Zachrisson, H. D., & Vandell, D. L. (2021). Adult outcomes of sustained high-quality early child care and education: Do they vary by family income? *Child Development*, *93*(2), 502–523. https://doi.org/10.1111/cdev.13696

Chzhen, Y., Gromada, A., & Rees, G. (2019). *Are the world's richest countries family friendly? Policy in the OECD and EU*. UNICEF Office of Research, Florence.

Clarke-Stewart, K. A. (1989). Infant day care: maligned or malignant? *The American Psychologist*, *44*(2), 266–273. https://doi.org/10.1037/0003-066X.44.2.266

Coley, R. L., Votruba-Drzal, E., Miller, P. L., & Koury, A. (2013). Timing, extent, and type of child care and children's behavioral functioning in kindergarten. *Developmental Psychology, 49*(10), 1859–1873. https://doi.org/10.1037/a0031251

Coley, R. L., Lombardi, C. M., & Sims, J. (2015). Long-Term implications of early education and care programs for Australian children. *Journal of Educational Psychology, 107*(1), 284–299. https://doi.org/doi:10.1037/a0037456

Côté, S. M., Boivin, M., Nagin, D. S., Japel, C., Xu, Q., Zoccolillo, M., Junger, M., & Tremblay, R. E. (2007). The role of maternal education and nonmaternal care services in the prevention of children's physical aggression problems. *Archives of General Psychiatry, 64*(11), 1305–1312. IF 13,747 https://doi.org/10.1001/archpsyc.64.11.1305

Côté, S. M., Mongeau, C., Japel, C., Xu, Q., Séguin, J. R., & Tremblay, R. E. (2013). Child care quality and cognitive development: Trajectories leading to better preacademic skills. *Child Development*, *84*(2), 752–766. https://doi.org/10.1111/cdev.12007

Crosby, D. A., Dowsett, C. J., Gennetian, L. A., & Huston, A. C. (2010). A tale of two methods: comparing regression and instrumental variables estimates of the effects of preschool child care type on the subsequent externalizing behavior of children in low-income families. *Developmental Psychology*, *46*(5), 1030–1048. https://doi.org/doi:10.1037/a0020384

Dearing, E., & Zachrisson, H. D. (2017). Concern over internal, external, and incidence validity in studies of child-care quantity and externalizing behavior problems. *Child Development Perspectives*, *11*(2), 133–138. https://doi.org/10.1111/cdep.12224

Dearing, E., Zachrisson, H. D., & Nærde, A. (2015). Age of entry into early childhood education and care as a predictor of aggression: faint and fading associations for young Norwegian children. *Psychological Science*, *26*(10), 1595–1607. https://doi.org/10.1177/095679761559501

Dowsett, C. J., Huston, A. C., Imes, A. E., & Gennetian, L. (2008). Structural and process features in three types of child care for children from high and low income families. *Early Childhood Research Quarterly, 23*, 69–93. https://doi.org/10.1016/j.ecresq.2007.06.003

Duncan, G. J., Engel, M., Claessens, A., & Dowsett, C. J. (2014). Replication and robustness in developmental research. *Developmental Psychology*, *50*(11), 2417–2425. https://doi.org/10.1037/a0037996

Duncan, G. J., Magnuson, K. A., & Ludwig, J. (2004). The endogeneity problem in developmental studies. *Research in Human Development*, *1*(1–2), 59–80. https://doi.org/10.1080/15427609.2004.9683330

Felfe, C., & Zierow, L. (2018). From dawn till dusk: Implications of full-day care for children's development. *Labour Economics*, *55*(December), 259–281. https://doi.org/10.1016/j.labeco.2018.10.006

Geoffroy, M., Séguin, J. R., Lacourse, E., Boivin, M., Tremblay, R. E., & Côté, S. (2012). Parental characteristics associated with childcare use during the first 4 years of life: Results from a representative cohort of Québec families. *Canadian Journal of Public Health*, *103*(1), 76–80. https://doi.org/10.1007/BF03404073

Goodman, Robert. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Huston, A. C., Bobbitt, K. C., & Bentley, A. (2015). Time spent in child care: How and why does it affect social development? *Developmental Psychology*, *51*(5), 621–634. https://doi.org/10.1037/a0038951

Jacob, J. I. (2009). The socio-emotional effects of non-maternal childcare on children in the USA: a critical review of recent studies. *Early Child Development and Care*, *179*(5), 559–570. https://doi.org/10.1080/03004430701292988

Kottelenberg, M. J, & Lehrer, S. F. (2017). Targeted or universal coverage? *Journal of Labor Economics*, *35*(3), 609–653. https://doi.org/10.1086/690652

Love, J.M., Harrison, L., Sagi-Schwartz, A., M.H, van Ij., Ross, C., J.A, U., … Chazan-Cohen, R. (2003). Child care quality matters: How conclusions may vary with context. *Child Development*, *74*(4), 1021–1033. https://doi.org/10.1111/1467-8624.00584

McCartney, K., Burchinal, M., Clarke-Stewart, A., Bub, K. L., Owen, M. T., Belsky, J., & the NICHD Early Child Care Research Network. (2010). Testing a series of causal propositions relating time in child care to children's externalizing behavior. *Developmental Psychology*, *46*(1), 1–17. https://doi.org/10.1037/a0017886

McCartney, K., & Rosenthal, S. (1991). Maternal employment should be studied within social ecologies. *Journal of Marriage and Family*, *53*(4), 1103–1107. https://doi.org/10.2307/353014

Melhuish, E. (2010). Why children, parents and home learning are important. In K. Sylva, E. Melhuish, P. Sammons, I. Siraj-Blatchford, & B. Taggart (Eds.), *Early childhood matters: Evidence from the effective pre-school and primary education project*. Routledge.

Morrissey, T. W. (2009). Multiple child-care arrangements and young children's behavioral outcomes. *Child Development*, *80*(1), 59–76. https://doi.org/10.1111/j.1467-8624.2008.01246.x

Morrissey, T. W. (2017). Child care and parent labor force participation: a review of the research literature. *Review of Economics of the Household, 15*(1), 1–24. https://doi.org/10.1007/s11150-016-9331-3

Nærde, A., Janson, H., & Ogden, T. (2014). BONDS (The behavior outlook Norwegian developmental study): A prospective longitudinal study of early development of social competence and behavior problems. (Report). Norsk Senter forstudier av problematferd og innovativ praksis. https://doi.org/10.13140/RG.2.2.10945.56169

Nærde, A., Ogden, T., Janson, H., & Zachrisson, H.D. (2014). Normative development of physical aggression from 8 to 26 months. *Developmental Psychology*, *50*(6), 1710–20. https://doi.org/10.1037/a0036324

NICHD Early Child Care Research Network. (1998). Early child care and self-control, compliance, and problem behavior at twenty-four and thirty-six months. *Child Development*, *69*(4), 1145–1170. https://doi.org/10.2307/1132367

NICHD Early Child Care Research Network. (2002a). Child-care structure→ process→ outcome: Direct and indirect effects of child-care quality on young children's

development. *Psychological science, 13*(3), 199–206. https://doi.org/10.1111/1467-9280.00438

NICHD Early Child Care Research Network. (2002b). Early child care and children's development prior to school entry: Results from the NICHD Study of Early Child Care. *American Educational Research Journal*, *39*(1), 133–164. https://doi.org/10.3102/00028312039001133

NICHD Early Child Care Research Network. (2003). Does amount of time spent in child care predict socioemotional adjustment during the transition to kindergarten? *Child Development*, *74*(4), 976–1005. https://doi.org/10.1111/1467-8624.00582

NICHD Early Child Care Research Network. (2004). Type of child care and children's development at 54 months. (2004). *Early Childhood Research Quarterly, 19*(2), 203–230. https://doi.org/10.1016/j.ecresq.2004.04.002

Organisation for Economic Co-operation Development, & Source. (2020). *Education at a Glance 2020: OECD Indicators*. https://doi.org/doi.org/10.1787/69096873-en

Organisation for Economic Co-operation Development, & Source. (2006). Starting Strong II: Early Childhood Education and Care. OECD Publishing.

Orri, M., Tremblay, R. E., Japel, C., Boivin, M., Vitaro, F., Losier, T., Brendgen, M. R., Falissard, B., Melchior, M. & Côté, S. M. (2019). Early childhood child care and disruptive behavior problems during adolescence: a 17-year population-based propensity score study. *Journal of Child Psychology and Psychiatry*, *60*(11), 1174–1182. https://doi.org/10.1111/jcpp.13065

Petitclerc, A., Côté, S., Doyle, O., Burchinal, M., Herba, C., Zachrisson, H. D., . . . Raat, H. (2017). Who uses early childhood education and care services? Comparing socioeconomic selection across five western policy contexts. *International Journal of Child Care and Education Policy* (Seoul), *11*(1), 1–24. https://doi.org/10.1186/s40723-

017-0028-8

Pingault, J. B., Tremblay, R. E., Vitaro, F., Japel, C., Boivin, M. and Côté, S. M. (2015). Early nonparental care and social behavior in elementary school: support for a social group adaptation hypothesis. *Child Development*, *86*, 1469–1488. https://doi.org/10.1111/cdev.12399

Ribeiro, L. A., & Zachrisson, H. D. (2019). Peer effects on aggressive behavior in Norwegian child care centers. *Child Development*, *90*(3), 876–893. https://doi.org/doi.org/10.1111/cdev.12953

Riley, R. D., Higgins, J. P. T., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ*, *342*(feb10 2), d549–d549. https://doi.org/10.1136/bmj.d549

Tremblay, R.E., Loeber, R., Gagnon, C., Charlebois, P., Larivee, S., & Leblanc, M. (1991). Disruptive boys with stable and unstable high fighting behavior patterns during junior elementary-school. *Journal of Abnormal Child Psychology*, *19*(3), 285–300. https://doi.org/10.1007/BF00911232

Vandell, D. L. (1996). Characteristics of infant child care: Factors contributing to positive caregiving: NICHD early child care research network. *Early Childhood Research Quarterly, 11*(3), 269–306. https://doi.org/10.1016/S0885-2006(96)90009-5

Veen, A., Veen, I., Heurter, A., Ledoux, G., Mulder, L., Paas,T., Leseman, P., Mulder, H., Verhagen, J., Slot, P. (2012). Pre-COOL cohortonderzoek. technisch rapport tweejarigencohort, eerste meting 2010-2011. [Pre-COOL cohortstudy. Technical report two-year-olds cohort, first measurement 2010-2011]. Kohnstamm Instituut.

Vernon-Feagans, L., & Cox, M. (2013). The Family Life Project: An epidemiological and developmental study of young children living in poor rural communities. *Monographs of the Society for Research in Child Development, 78*(5), 1–150. https://doi.org/10.1111/mono.12046

Votruba-Drzal, Coley, R. L., & Chase-Lansdale, L. P. (2004). Child care and low-income children's development: Direct and moderated effects. *Child Development*, *75*(1), 296–312. https://doi.org/10.1111/j.1467-8624.2004.00670.x

Weinert, S., Roßbach, H.-G., Faust, G., Blossfeld, H.-P. & Artelt, C. (2013). Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter (BiKS-3-10) (Version 6) [Educational processes, competence development and selection decisions in pre-school and school age]. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. http://doi.org/10.5159/IQB_BIKS_3_10_v6

World Bank, International Comparison Program database. (2011). *GDP per capita, PPP (constant 2011 international $)*. https://databank.worldbank.org/metadataglossary/jobs/series/NY.GDP.PCAP.PP.KD

World Bank, Poverty and Equity database. (n.d.). *Gini coefficient (income after taxes and transfers)*. https://povertydata.worldbank.org/poverty/home/

World Bank, Education Statistics (EdStats). (n.d.). *Government expenditure on pre-primary education as % of GDP (%) and Pupil-teacher ratio, preprimary*. https://datatopics.worldbank.org/education/indicators

Yamauchi, C., & Leigh, A. (2011). Which children benefit from non-parental care? *Economics of Education Review*, *30*(6), 1468–1490. https://doi.org/10.1016/j.econedurev.2011.07.012

Zachrisson, H. D., Dearing, E., Lekhal, R., & Toppelberg, C. O. (2013). Little evidence that time in child care causes externalizing problems during early childhood in Norway. *Child Development*, *84*(4), 1152–1170. https://doi.org/10.1111/cdev.12040

Zachrisson, H. D., Owen, M. T., Nordahl, K. B., Ribeiro, L., & Dearing, E. (2021). Too early for early education? effects on parenting for mothers and fathers. *Journal of Marriage and Family, 83*(3), 683–698. https://doi.org/10.1111/jomf.12701

Do More Hours in Center-based Care Cause More Externalizing Problems?

A Cross-National Replication Study

Supplementary material

Table S1

Standard Deviation Changes in Externalizing Behaviors as a Function of 10hr Increase in Time in Nonparental Care

	EMIGARDE	**QLSCD**	**Pre-COOL**	**FLP**	**NICHD-SECCYD**
	parent-report	**parent-report**	**teacher-report**	**parent-report**	**parent-report**
	***b* (*SE*)**	***b* (*SE*)**	***b* (*SE*)**	***b* (*SE*)**	***b* (*SE*)**
Random effects					
Quantity	.010 (.022)	.011 (.007)	.122*** (.036)	.007 (.010)	.021 (.059)
Nonlinear specification					
Quantity	.020 (.023)	.012+ (.007)	.122** (.036)	.001 (.010)	.019 (.061)
Quantity- squared	.010 (.010)	.004 (.004)	.005 (0.02)	-.003 (.003)	.028 (.028)
Interaction with education					
Quantity	.009 (.060)	-.005 (.013)	.095 (.059)	.010 (.008)	.012 (.067)
Education	-.447** (.153)	.048 (.059)	-.151* (.086)	-.291*** (.061)	-.226** (.059)
Education x Quantity	.001 (.062)	.022 (.014)	0.033 (.051)	.029 (.019)	.033 (.025)
Interaction with income					
Quantity	-.053 (.055)	.014+ (.008)	.105 (.072)	.009 (.009)	.013 (.070)

Income	-.015 (.028)	.091* (.046)	-.020 (.014)	-.100*** (.015)	-.001 (.010)
Income x Quantity	.009 (.008)	-.013 (.015)	.003 (.007)	.000 (.006)	.003 (.005)
Fixed effects					
Quantity	.034 (.048)	-.002 (.014)	-.002 (.060)	.004 (.011)	-.010 (.030)
Nonlinear specification					
Quantity	.044 (.068)	.005 (.008)	-.006 (.063)	.006 (.011)	-.011 (.034)
Quantity- squared	-.006 (.027)	.003 (.005)	.007 (.028)	-.002 (.003)	.001 (.014)
Interaction with education					
Quantity	.060 (.167)	.040 (.030)	.009 (.107)	.555 (.800)	-.025 (.049)
Education x Quantity	-.028 (.173)	-.059+ (.034)	.008 (.104)	.027 (.019)	.024 (.057)
Interaction with income					
Quantity	-.023 (.167)	.003 (.015)	-.016 (.101)	.004 (.011)	-.033 (.040)
Income x Quantity	.008 (.022)	-.030 (.040)	.002 (.014)	.001 (.006)	.006 (.008)

Notes. Coefficients for hours in care are reported with 10-hour increments. Fixed effects models controlled for number of siblings, single parent, maternal employment status. Random effects models controlled for maternal employment status, number of siblings, single parent, maternal education (at timepoint 1), parents' migrant status (or ethnicity in the case of the U.S.), sex of the child, and family income. + $p<.10$, * $p<.05$, ** $p<.01$, ***$p<.001$.

Table S2

Descriptive Statistics for the Covariates in the Analytic Samples

	EMIGARDE		**QLSCD**		**BIKS**		**Pre-COOL**		**BONDS**		**FLP**		**NICHD SECCYD**	
Variables	*M* (*SD*)	min, max	*M* (*SD*)	min, max	*M* (*SD*)	min, max	*M* (*SD*)	min, max	*M* (*SD*)	min, max	*M* (*SD*)	min, max	*M* (*SD*)	min, max
Covariates														
Time-variant covariates														

Maternal employ T1[a]	57%	0, 1	70%	0, 1	51%	0, 1	72%	0, 1	21.11 (16.57)	0, 80	62%		22.21 (19.09)	0, 105
Maternal employ T2	50%	0, 1	74%	0, 1	55%	0, 1	82%	0, 1	19.53 (16.97)	0, 60	63%		22.44 (19.17)	0, 80
Maternal employ T3	65%	0, 1	77%	0, 1	--	--	--	--	25.1 (15.98)	0, 60	64%		23.14 (19.14)	0, 98
Maternal employ T4	--	--	75%	0, 1	--	--	--	--	--	--			--	--
Single T1	10%	0, 1	9%	0, 1	8%	0, 1	8%	0, 1	8%	0, 1	37%		14%	0, 1
Single T2	11%	0, 1	12%	0, 1	9%	0, 1	8%	0, 1	9%	0, 1	35%		15%	0, 1
Single T3	11%	0, 1	12%	0, 1	--	--	--	--	9%	0, 1	33%		17%	0, 1
Single T4	--	--	13%	0, 1	--	--	--	--	--	--			--	--
Siblings/New sib T1	53%	0, 1	0.84 (0.89)	0, 4	1.11 (0.9)	0, 6	0.79 (0.83)	0, 4	14%	0, 1	1.38 (1.19)	0, 8	2.01 (1.05)	1, 10
Siblings/New sib T2	70%	0, 1	0.97 (0.87)	0, 4	1.16 (0.85)	0, 6	1.32 (0.96)	0, 4	11%	0, 1	1.47 (1.20)	0, 8	2.13 (1)	1, 8
Siblings/New sib T3	77%	0, 1	1.11 (0.85)	0, 4	--	--	--	--	18%	0, 1	1.54 (1.17)	0, 8	2.27 (0.98)	1, 8
Siblings/New sib T4	--	--	1.18 (0.86)	0, 4	--	--	--	--	--	--			--	--
Time-invariant covariates														
Maternal education (more than high school)	91%	0, 1	4.17 (2.18)	1, 7	25%	0, 1	80%	0, 1	58%	0, 1	81%		69%	0, 1
Gender (boy)	50%	0, 1	51%	0, 1	52%	0, 1	51%	0, 1	52%	0, 1	49%		52%	0, 1
Family income[b]	7.28 (2.11)	1, 9	21%	0, 1	1452.318 (525.37)	250, 3076.9	7.03 (2.90)	1, 17	11%	0, 1	1.93 (1.62)	0, 13.78	52.37 (41.64)	25, 400
Immigrant status														
Any immigrant status	21%	0, 1	--	--	26%	0, 1			--	--			--	--
European immigrant	--	--	0.03 (0.17)	0, 1	--	--			--	--			--	--
Non-European immigrant	--	--	0.09 (0.28)	0, 1	--	--			--	--			--	--
Western immigrant	--	--	--	--	--	--	0.07 (0.26)	0, 1	7%	0, 1			--	--
Non-Western immigrant	--	--	--	--	--	--	0.17 (0.38)	0, 1	6%	0, 1			--	--
Ethnicity														

White	--	--	--	--	--	--			--	--	57%		80%	0, 1
African-American	--	--	--	--	--	--			--	--	42%		13%	0, 1
Latinx[(c)]	--	--	--	--	--	--			--	--	1.2%		5%	0, 1
Other											1%		2%	0, 1
Main predictor for secondary analyses														
Quantity of nonparental care (hours)														
Nonparental T1	29.72 (17.42)	0, 52	19.02 (18.6)	0, 50	--	--	14.87 (9.6)	0, 50	--	--	23.08 (21.70)	0, 117	20.79 (19.4)	0, 50
Nonparental T2	30.24 (15.76)	0, 50	21.52 (18.9)	0, 50	--	--	12.77 (7.90)	0, 50	--	--	24.12 (21.40)	0, 116	22.11 (19.14)	0, 50
Nonparental T3	30.88 (17.44)	0, 55	25.84 (17.82)	0, 50	--	--	--	--	--	--	23.66 (21.60)	0, 115	25.91 (17.3)	0, 53
Nonparental T4			26.03 (17.95)	0, 50	--	--	--	--	--	--	--	--	--	--
Nonparental *within SD*	(12.09)	--	(10.50)	--	--	--	(3.49)	--	--	--	(16.17)	--	(10.61)	--

Notes. (a) Self-reported of hours per week in BONDS and NICHD SECCYD. (b) Income to needs ratio (INR) in FLP, total family income in thousands of dollars in NICHD SECCYD. (c) Lantix doesn't count toward the 100% in the FLP data, it was considered ethnicity (i.e., in addition to race).

Cumulative hours in center-based care specification

While cumulative hours models are not our primary models of interest, these models offer an important alternative specification that address the possibility that it is sustained time in a large number of hours that causes risk. The cumulative hour models estimate the within-person impact of an accumulation of hours in care which occurs through sustained time in care. The following example offers a short explanation of the cumulative hours specification. Consider three children whose hours increase at a similar rate across three observations. In the first table, hours

are coded as the number of hours the child is in care at that time point. In the second table, cumulative hours for these same three children are calculated at each time point.

In the first case, the size of changes in hours would be identical for these three children:

Child	Time 1	Time 2	Time 3	**T2-T1**	**T3-T1**
1	0	10	20	**10**	**20**
2	10	20	30	**10**	**20**
3	20	30	40	**10**	**20**

In the second case, children who accumulate more hours in care would have larger increases in hours:

Child	Time 1	Time 2	Time 3	**T2-T1**	**T3-T1**
1	0	10	30	**10**	**30**
2	10	30	60	**20**	**50**
3	20	50	90	**30**	**70**

Another distinction between these two approaches is which children drop out due to stability.

In the first case, children who were never in care and children in stable hours of care drop out of the model. In the second case, children never in care and those who are only in care at T1 drop out.

Table S3

Standard Deviation Changes in Externalizing Behaviors as a Function of 10hr Increase in Time in Cumulative Center-Based Care

	EMIGARDE	**QLSCD**	**BIKS**	**Pre-COOL**	**BONDS**	**FLP**	**NICHD SECCYD**
	b (SE)	***b (SE)***	***b (SE)***	***b (SE)***	***b (SE)***	***b (SE)***	***b (SE)***
Random effects							
Quantity	.007 (.011)	-.008 (.006)	.025 (.040)	.094*** (.023)	.003* (.001)	.000 (.009)	.001 (.014)
Interaction with maternal education							
Quantity	-.016 (.029)	-.021^{+} (.011)	.034 (.042)	.077^{+} (.041)	.003* (.001)	-.008 (.006)	.002 (.020)
Education	-.530** (.190)	.018 (.011)	-.25 (.138)	-.156^{+} (.084)	.030 (.150)	-.308*** (.061)	-.228** (.070)
Education x Quant	.024 (.030)	.017 (.011)	-.023 (.032)	.021 (.036)	-.000 (.001)	-.015 (.014)	-.001 (.015)
Interaction with family income							
Quantity	-.042 (.030)	-.006 (.006)	.025 (.040)	.086 (.052)	.003* (.001)	.007 (.006)	.012 (.021)
Income	-.051^{+} (.027)	.180^{+} (.098)	-.001 (.016)	-.022 (.013)	-.099 (.144)	-.100*** (.015)	.009 (.014)
Income x Quant	.007^{+} (.004)	-.012 (.012)	.001 (.003)	.001 (.006)	.003 (.002)	.002 (.003)	-.003 (.002)

Fixed effects							
Quantity	-.024 (.046)	-.025* (.011)	-.064 (.075)	-.031 (.079)	.003 (.002)	.-003 (.009)	-.026 (.028)
Interaction with maternal education							
Quantity	-.057 (.063)	-.038* (.016)	-.057 (.077)	-.027 (.134)	.002 (.004)	-.002 (.010)	-.019 (.035)
Education x Quant	.035 (.043)	.015 (.013)	-.023 (.035)	-.005 (.123)	.000 (.000)	-.015 (.014)	-.010 (.026)
Interaction with family income							
Quantity	-.076 (.058)	-.024* (.012)	-.062 (.075)	-.088 (.132)	.002 (.003)	-.003 (.009)	-.012 (.033)
Income x Quant	.007 (.005)	-.013 (.013)	.001 (.004)	.008 (.013)	.002 (.002)	.003 (.003)	-.003 (.004)

Notes. Coefficients for hours in care are reported with 10-hour increments. Fixed effects models controlled for number of siblings, single parent, maternal employment status. Random effects models controlled for maternal employment status, number of siblings, single parent, maternal education (at timepoint 1), parents' migrant status (or ethnicity in the case of the U.S.), sex of the child, and family income. + $p<.10$, * $p<.05$, ** $p<.01$, *** $p<.001$.

Table S4

Standard Deviation Changes in Externalizing Behaviors as a Function of Transitions Between No Care And Part-Time Care to Full-Time Care

	EMIGARDE	**QLSCD**	**BIKS**	**Pre-COOL**	**BONDS**	**FLP**	**NICHD SECCYD**
	b (SE)	***b (SE)***	***b (SE)***	***b (SE)***	***b (SE)***	***b (SE)***	***b (SE)***
Random effects							
No care	-.017 (.081)	-.035 (.028)	.043 (.239)	-.443*** (.099)	-.070 (.090)	-.039 (.042)	.131 (.126)
Part-time	-.043 (.091)	-.011 (.040)	-.191* (.080)	-.262** (.078)	-.115^{+} (.061)	.005 (.050)	-.120 (.072)
Fixed effects							
No care	.021 (.107)	-.372(.477)	.276 (.639)	-.215 (.182)	-.072 (.334)	-.040 (.042)	.321^{+} (.176)
Part-time	-.016 (.102)	.017 (.054)	-.211 (.128)	-.106 (.130)	-.048 (.088)	-.001 (.051)	.024 (.076)

Notes. Coefficients for hours in care are reported with 10-hour increments. Fixed effects models controlled for number of siblings, single parent, maternal employment status. Random effects models controlled for maternal employment status, number of siblings, single parent, maternal education (at timepoint 1), parents' migrant status (or ethnicity in the case of the U.S.), sex of the child, and family income. + $p<.10$, * $p<.05$, ** $p<.01$, *** $p<.001$.

Table S5

Standard Deviation Changes in Externalizing Behaviors as a Function of 10hr Increase in Time in Center-Based Care Without Controlling for Maternal Employment and With Time-varying Income

	EMIGARDE		**QLSCD**		**BIKS[(a)]**	**Pre-COOL[(a)]**	**BONDS[(a)]**	**FLP**		**NICHD SECCYD**	
	With time-varying income	***Without maternal employ-ment***	***With time-varying income***	***Without maternal employ-ment***	***Without maternal employ-ment***	***Without maternal employ-ment***	***Without maternal employ-ment***	***With time-varying income***	***Without maternal employ-ment***	***With time-varying income***	***Without maternal employ-ment***
Random effects											
Quantity	.004 (.020)	.013 (.020)	.007 (.008)	.008 (.007)	.089 (.055)	.136*** (.034)	.037^{+} (.019)	.006 (.011)	.005 (.010)	-.040 (.031)	-.031 (.033)
Interaction with maternal education											
Quantity	.008 (.064)	.017 (.065)	-.002 (.015)	.000 (.014)	.133^{+} (.069)	.115^{+} (.058)	.059^{+} (.031)	.010 (.009)	.008 (.009)	-.051 (.044)	-.043 (.047)
Education	-.448** (.154)	-.436** (.155)	.030 (.058)	.034 (.058)	-.122 (.080)	-.163^{+} (.083)	.004 (.101)	-.101 (.091)	-.100 (.092)	-.215*** (.058)	-.223** (.062)
Education x Quant	-.005 (.067)	-.005 (.067)	.012 (.016)	.011 (.016)	-.109 (.104)	.026 (.055)	-.045 (.044)	.004 (.026)	.004 (.026)	.016 (.030)	.019 (.030)
Interaction with family income											
Quantity	-.047 (.057)	-.060 (.061)	.005 (.007)	.006 (.008)	.088 (.055)	.098 (.071)	.033^{+} (.020)	.007 (.009)	.006 (.009)	-.045 (.051)	-.037 (.052)
Income	-.014 (.020)	-.013 (.020)	.069 (.052)	.085+ (.046)	.001 (.009)	-.027* (.012)	.072 (.072)	-.003(.017)	-.008 (.017)	-.005 (.011)	.005 (.011)

Income x Quant	.007 (.008)	.010 (.009)	.009 (.017)	.009 (.017)	-.001 (.009)	.006 (.008)	.038 (.069)	-.001 (.010)	-.001 (.010)	.001 (.006)	.002 (.006)
Nonlinear specification											
Quantity	.009 (.023)	.019 (.23)	-.003 (.010)	-.002 (.009)	.093 (.056)	.142*** (.033)	.071* (.032)	.000 (.014)	-.002 (.014)	-.082 (.059)	-.078 (.061)
Quantity-squared	.006 (.015)	.007 (.015)	.009 (.006)	.010+ (.006)	.023 (.041)	.023 (.019)	.019 (.016)	.003 (.005)	.005 (.005)	.029 (.021)	.032 (.021)
Fixed effects											
Quantity	.029 (.059)	.044 (.057)	-.016 (.027)	-.012 (.027)	.063 (.108)	.065 (.109)	.007 (.046)	.001 (.011)	.001 (.011)	-.050 (.053)	-.046 (.052)
Interaction with maternal education											
Quantity	-.127 (.186)	-.116 (.184)	.002 (.064)	.004 (.065)	.131 (.128)	.064 (.157)	.018 (.071)	.000 (.011)	.000 (.011)	-.046 (.092)	-.040 (.091)
Education	--	--	--	--	--	--	--	-.063 (.096)	-.063 (.096)	--	--
Education x Quant	.173 (.196)	.178 (.196)	-.023 (.068)	- .020 (.069)	-.181 (.190)	.001 (.136)	-.020 (.087)	.003 (.027)	.003 (.027)	-.005 (.093)	-.008 (.094)
Interaction with family income											
Quantity	-.133 (.201)	-.110 (.209)	-.020 (.028)	.021 (.028)	.074 (.105)	.030 (.162)	.019 (.049)	.001 (.011)	.002 (.011)	-.033 (.064)	-.083 (.073)
Income	-.047 (.055)	--	-.092 (.153)		--	--	--	-.002 (.017)	-.006 (.017)	-.012 (.021)	--
Income x Quant	.022 (.026)	.021 (.028)	.025 (.065)	.005 (.074)	.009 (.017)	.005 (.022)	-.090 (.132)	.001 (.010)	.001 (.010)	-.004 (.009)	.009 (.013)
Nonlinear specification											
Quantity	.109 (.098)	.126 (.097)	.044 (.073)	.047 (.073)	.063 (.109)	.067 (.115)	-.009 (.049)	-.002 (.014)	-.004 (.014)	-.085 (.087)	-.082 (.087)

Quantity-squared	-.043 (.041)	-.044 (.041)	-.018 (.022)	-.017 (.021)	.004 (.080)	-.003 (.047)	-.018 (.028)	.002 (.006)	.003 (.006)	.012 (.029)	.012 (.029)

Note. (a) Time-varying education/income was not available. Coefficients for hours in care are reported with 10-hour increments. Fixed effects models controlled for number of siblings, single parent, maternal employment status. Random effects models controlled for maternal employment status, number of siblings, single parent, maternal education (at timepoint 1), parents' migrant status (or ethnicity in the case of the U.S.), sex of the child, and family income. + $p<.10$, * $p<.05$, ** $p<.01$, ***$p<.001$.

Table S6

Logit Models of Differences Between Children With Stable Hours Compared With Children who Were not Stable (Center-Based Care and Nonparental Care)

Variables	EMIGARDE		QLSCD		BIKS	Pre-COOL		BONDS	FLP		NICHD SECCYD	
	Center care	Any non-parental care	Center care	Any non-parental care	Center care	Center care	Any non-parental care	Center care	Center care	Any non-parental care	Center care	Any non-parental care
% stable	11.15	7.13	34.36	11.87	28.32			9.03	22.5	37	20.9	9.61
Maternal education	0.037 (.095)	-.032 (.120)	.061* (.024)	.015 (.038)	.196 (.201)	.299 (.388)	.230 (.382)	-.027 (.046)	-.004 (.372)	-.005 (.313)	.544*** (.153)	.304 (.211)
Immigrant category 1	[omitted]	[omitted]	.058 (.279)	-.385 (.379)		-.180 (.553)	-.159 (.530)	-.089 (.408)	.506* (.232)	.480* (.208)	[omitted]	[omitted]
Immigrant category 2	-.134 (.362)	-.163 (.441)	.923*** (.196)	.595* (.267)	.198 (.227)	-1.021* (.507)	-.757 (.485)	.054 (.457)	[omitted]	[omitted]	.665 (.528)	1.116^{+} (.633)
Immigrant category 3	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	.176 (.490)	.543 (.566)
Immigrant category 4	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	.413 (.585)	1.199 (.767)

Boy	.155 (.293)	-.477 (.370)	.036 (.093)	-.256^{+} (.145)	-.098 (.192)	-.055 (.239)	.079 (.233)	.255 (217)	-.044 (.202)	-.132 (.184)	.134 (.140)	208 (.190)
Siblings	-.275 (.200)	-0.436^{+} (240)	-.172 ** (.060)	-.227** (.082)	-.028 (.110)	.030 (.157)	.078 (.153)	1.008 (.680)	.020 (.098)	-.093 (.091)	-.006 (.075)	-.100 (.097)
Single parents	-1.025 (.729)	-2.053^{+} (1.065)	.673** (.220)	.767* (.316)	1.123* (.460)	.115 (.593)	.432 (.571)	.189 (.516)	.428 (.273)	.859** (.253)	.138 (.240)	0.029 (.321)
Maternal employment	.726^{+} (.390)	.741 (.487)	1.287*** (.164)	2.537*** (.206)	.073 (.206)	.492 (.328)	.222 (.324)	-.027** (.009)	1.954*** (.395)	2.175*** (.32)	-.019*** (.005)	-.006 (006)
Income	-.016 (.103)	.105 (.123)	.462* (.199)	.066 (.246)	n.a.	.031 (.046)	.057 (.045)	.273 (.397)	.229*** (.085)	.238** (.084)	.255*** (.048)	.190** (.062)

Note. 0=stable hours,1=not stable. All time-varying covariates are averaged across timepoints. Children with stable zero hours in care are excluded. Coefficients for hours in care are reported with 10-hour increments. Fixed effects models controlled for number of siblings, single parent, maternal employment status. Random effects models controlled for maternal employment status, number of siblings, single parent, maternal education (at timepoint 1), parents' migrant status (or ethnicity in the case of the U.S.), sex of the child, and family income. + p<.10, * p<.05, ** p<.01, ***p<.001.