Article

The psychosocial health of children born after medically assisted reproduction: Evidence from the UK Millennium Cohort Study☆

Anna Barbuscia[a,*], Mikko Myrskylä[a,b], Alice Goisis[a]

[a] LSE Social Policy Department, MPIDR, United Kingdom
[b] Department of Social Research, Helsinki University, Finland

ARTICLE INFO

ABSTRACT

The increasing number of children conceived through medically assisted reproduction (MAR, including IVF/ICSI, intrauterine insemination and ovulation induction) has led to concerns about the potential negative effects of fertility treatments on children's psychosocial health. Some studies suggest that MAR children might be at higher risk of developing psychosocial problems when they enter adolescence. However, very few studies have examined the development of MAR children after childhood. Moreover, even though parental socio-economic characteristics are known to be highly correlated with children's psychosocial development, most existing studies on the outcomes of MAR children did not take into account the selective characteristics of the couples who accessed fertility treatments. Using data from waves 1–6 of the UK Millennium Cohort Study, we compare the psychosocial health, as measured by the Strengths and Difficulties Questionnaire, of MAR children to that of naturally-conceived (NC) children, up to and including the age of 14. We control for a wide range of time-constant child and parental characteristics that might confound the association between MAR and the psychosocial health of children. Results from multilevel random intercept models that do not account for parental characteristics show that MAR children have a lower incidence of psychosocial problems than NC children. In models that control for parental characteristics, MAR children are found to have a higher incidence of psychosocial problems than NC children at age three, which suggests that high parental resource levels both explain the advantage of MAR children in unadjusted models, and mask the potentially adverse effects of MAR at young ages. However, in the fully adjusted models in which MAR children have more psychosocial problems at young age, the differences with respect to NC children decrease with age and become statistically and substantively negligible by end of follow-up at age 14. This result suggests that the use of MAR does not increase children's risk of having psychosocial problems at the onset of adolescence.

1. Introduction

The number of children born through medically assisted reproduction (MAR) – which includes a range of techniques, such as ovulation induction, artificial insemination, in vitro fertilisation (IVF), and intracytoplasmic sperm injection (ICSI) – has increased markedly over the last few decades. Since 1978, when the first baby was born *in vitro*, an estimated 5.4 million babies conceived through MAR have been born worldwide (Calhaz-Jorge et al., 2016). This has, caused some scholars to raise concerns about the wellbeing and development of these children, including about a higher risk of developing psychosocial problems (Colpin, 2002; Hart & Norman, 2013a, 2013b; Präg & Mills, 2015; van Balen, 1998).

Various explanations for why children conceived through MAR might experience more psychosocial health problems than naturally conceived (NC) children have been proposed. First, the use of MAR is associated with increased risks of adverse birth outcomes, such as low birth weight and prematurity (Hart & Norman, 2013a; Ludwig, Sutcliffe, Diedrich, & Ludwig, 2006; van Balen, 1998), which might in turn be associated with worse outcomes later in life (Boardman, Powers, & Hummer, 2002). Second, compared to natural conception, conception through MAR might be associated with higher levels of parental stress, which could influence parent-child relationships and children's psychosocial development (Bernstein, 1990; Colpin & Soenen, 2002; Golombok, Cook, Bish, & Murray, 1995; Wagenaar et al., 2008). Scholars have argued that a child born after MAR might be seen as especially precious, which could lead the parents to be overprotective or to have exaggerated expectations (Wagenaar et al., 2009).

Previous studies on the psychosocial health of MAR children have provided mixed results. While no overall differences have been found in the incidence of problems during childhood (Colpin, 2002; Golombok et al., 1996; Hart & Norman, 2013b), there is some evidence of increased risks (of depression) in early adolescence (Wagenaar et al., 2008). Scholars have argued that the mode of conception might be especially likely to affect the parent-child relationship during the stage of the life course when children become more independent and develop their identity (Colpin & Soenen, 2002; Golombok, MacCallum, & Goodman, 2001; Golombok et al., 2002; Hart & Norman, 2013b; Owen & Golombok, 2009). Thus, early adolescence could be an especially arduous time for MAR children. However, these findings are not conclusive, and only a few studies have examined the psychosocial outcomes of MAR children after childhood (Hart & Norman, 2013b; Ilioi & Golombok, 2015; Ludwig et al., 2006). Furthermore, most of the existing studies on this topic have examined the outcomes in different age groups in isolation, and have rarely scrutinised how differences between MAR and naturally conceived children develop over time (examples of studies following MAR families over time are Colpin & Bossaert, 2008; Golombok, Ilioi, Blake, Roman, & Jadva, 2017; Golombok et al., 2002).

As research on the development of non-cognitive skills has underlined the importance of looking at the progression of the acquisition of such skills across the stages of childhood, rather than at a single point in time (Cunha, Heckman, & Schennach, 2010; Heckman, 2006), it is crucial that we analyse the longer-term psychosocial health of MAR children and explore how it differs from that of NC children as they grow older. It is also important that we take into account more explicitly the relatively high socio-economic status of the parents who have conceived children through MAR. Since in most countries the costs of MAR treatments are not covered by public insurance (Chambers et al., 2014; Chandra, Copen, & Stephen, 2014), the couples who have access to these treatments tend to be highly educated and affluent; and thus have characteristics that are positively associated with children's cognitive and non-cognitive development (Ermisch, 2008; Feinstein, 2002; Heckman, 2006).

Using data from six waves of the UK Millennium Cohort Study (MCS), we compare levels of psychosocial problems of MAR children with NC children up to the age of 14. We use parent-reported measures from the Strengths and Difficulties Questionnaire (SDQ), and consider both the total difficulties scores and the internalising and externalising problems sub-scales. We begin our analysis by looking at unadjusted associations. We then exploit the richness of the MCS data to assess whether and, if so, how the unadjusted associations change after we adjust for the characteristics of the children and the parents that might confound or mediate the association between the use of MAR and the psychosocial health of children born after MAR treatments.

2. The psychosocial development of children born after fertility treatments: what is already known

There is consistent evidence that the risk of poor birth outcomes is higher for MAR children than for NC children (see Hart & Norman, 2013a; Pinborg, Henningsen, Malhau, & Loft, 2013 for a review); primarily because of the high prevalence after fertility treatments of multiple births, which are associated with increased risks of low birth weight and prematurity (Ludwig et al., 2006; Pison, Monden, & Smits, 2015). However, the studies that examined the physical health and cognitive development of children at older ages have found that, on average, MAR and NC children have similar outcomes (Hart & Norman, 2013a). A few studies have found that MAR children have better cognitive skills than NC children up to the age of seven, and have suggested that these differences are mainly attributable to the advantaged socio-economic characteristics of the couples who access fertility treatments (Barbuscia & Mills, 2017; Carson et al., 2011).

While studies on parent-child relationships and children's

psychosocial health have provided mixed results, most have reported similar outcomes for MAR and NC families. There is some evidence that having an MAR child has positive effects on parenting, with these parents exhibiting higher levels of emotional involvement, lower levels of stress, and a greater sense of competence (Golombok et al., 2001; Hahn & DiPietro, 2001). The results of the European Study of Assisted Reproduction Families (the first phase refers to when children were 4–8 years old and follow ups to ages 11–12 and 18), suggested that positive relationships between mothers and children is a feature of MAR throughout childhood (Golombok et al., 1996, 2002) and the adolescent years (Owen & Golombok, 2009).

Most of the existing studies that compared the incidence of behavioural and socio-emotional problems in MAR and NC children found no significant differences between these groups up to age eight (Colpin 2002; Golombok et al., 1995; Hart & Norman, 2013b; Wagenaar et al., 2009), and 11–12 (Golombok et al., 2002). However, Hahn and DiPietro (2001) found that MAR children had fewer emotional and behavioural problems according to their teachers; while other studies found that MAR children displayed more problematic behaviours according to their mothers (Gibson et al., 1998) and their teachers (Levy-Shiff, Vakil, & Dimitrovsky, 1998). Using data from the MCS, Carson et al. (2013), found that at the ages of five and seven, children born after assisted reproductive technology (i.e., through IVF or ICSI) had more psychosocial problems than children who were born after a naturally conceived and planned pregnancy.

While the overall evidence suggests that the psychosocial health of MAR children does not differ from that of NC children during childhood, some studies have pointed to an increased risk of experiencing problems in adolescence (Hart & Norman, 2013b; Wilson, Fisher, Hammarberg, Amor, & Halliday, 2011). For example, Wagenaar et al. (2009) found that IVF-conceived adolescents were more likely than their naturally conceived peers to score in the borderline/clinical range for depression. However, a follow-up analysis based on a self-reported questionnaire found no significant differences in the depression levels of the two groups of children at ages ranging from 11 to 18 (Wagenaar et al., 2011). Other studies did not find any differences in adolescents' psychological well-being at age 14 (Golombok et al., 2017), 15–16 (Colpin & Bossaert, 2008) and in the parent-reported and self-reported emotional, hyperactivity, peer problems, and prosocial behaviour scores of MAR and NC children up to age 21 (Zhu et al., 2011). Bay, Mortensen, Hvidtjorn, and Schioler (2013) found a systematically small increased risk of mental disorders in children born after ovulation/intrauterine insemination (but not after IVF/ICSI) compared with NC children aged between 7 and 18 years. In sum, the existing evidence on the longer-term psychosocial health of MAR children is inconclusive, as the handful of studies that have looked at the outcomes of these children beyond the childhood period have reported mixed findings (Bay et al., 2013; Ludwig et al., 2006; Wagenaar et al., 2011; Hart & Norman, 2013a). Because psychosocial problems and symptoms of clinical depression often do not emerge prior to the onset of adolescence (Patalay & Fitzsimons, 2017), findings on the psychosocial health of MAR children during childhood tell us little about the prevalence of such conditions during the adolescent period.

It is difficult to establish a priori whether MAR children are more likely than their naturally conceived peers to suffer from psychosocial problems during adolescence. On the one hand, scholars have suggested that MAR children might be at higher risk than NC children of experiencing problems as they enter adolescence and start to develop their identity (Golombok et al., 1995, 2017). It has also been suggested that the parents of a MAR child might find it especially difficult to cope with their child growing up and becoming more autonomous and independent (Colpin & Soenen, 2002; Golombok et al., 2001), which could in turn have negative consequences for the child's wellbeing. In addition, parents who used donor conception often do not tell their MAR children how they were conceived until the children have reached adolescence (Golombok et al., 1996; Golombok, Blake, Casey, Roman,

& Jadva, 2013; Ilioi & Golombok, 2015). This news may have detrimental effects on the wellbeing of these children, as they might feel confused about their identity or deceived by their parents. For example, one study found that MAR children who had been told how they were conceived exhibited more psychosocial problems than their non-informed counterparts, although they were still in the normal range (Colpin & Soenen, 2012). There is also evidence that when MAR children discover how they were conceived, their relationship with their parents and their psychosocial health may be negatively affected (Golombok et al., 2001).

On the other hand, it is important to take into account that most MAR children come from highly advantaged families. The literature has consistently shown that children's development of cognitive and non-cognitive skills (such as self-control, social development and behaviour problems) depends on their early family environment and the resources and investments of their parents (Conti & Heckman, 2014; Cunha & Heckman, 2007; Cunha et al., 2010; Heckman, 2006). Several studies have reported that children from more advantaged families are at lower risk of developing psychosocial problems than their less advantaged peers (Lewis, Hope, & Pearce, 2014; Patalay & Fitzsimons, 2017). Compared to parents who conceive naturally, parents who access (expensive) fertility treatments tend to be older, better educated, and have higher earning (Chambers et al., 2014; Chandra et al., 2014). These characteristics are likely to be associated positively with the psychosocial health of MAR children. As the advantages of couples who conceive through MAR might offset the greater likelihood that they will experience adverse birth outcomes or difficult parent-child relationships, the overall psychosocial development of MAR children could turn out to be positive.

In light of the reviewed literature, we believe there is a need for more research on the longer-term psychosocial health of MAR children, including during the period after childhood. It is also crucial to consider the existence of socio-economic gradients in children's development, and how this pattern might affect the outcomes of MAR children, who represent a selected and advantaged sub-group in terms of the demographic and socio-economic characteristics of their parents.

3. Data and methods

The UK Millennium Cohort Study (MCS) is a longitudinal survey that follows around 19,000 children born in the UK in 2000–2002 and their families. The sample is clustered geographically and stratified to over-represent areas of England with relatively high proportions of ethnic minorities and high levels of child poverty, as well as areas in the three smaller countries of Wales, Scotland, and Northern Ireland. Baseline interviews were conducted when the children were approximately nine months old, and follow-up interviews were conducted when the children were around three, five, seven, 11, and 14 years old. At the baseline interview, respondents (the cohort member's mother in 99% of cases) were asked whether they had used any fertility treatment to conceive. The MCS includes detailed information about the demographic, health, and socio-economic characteristics of the respondents and their families, including data on their children's development based on several measures of non-cognitive skills.

Our sample includes all of the children who were present at the first wave, and for whom information on psychosocial health is available at age three (the first time it was collected in the MCS) and in at least one other wave. Our final sample includes 12,989 children, of whom 404 were conceived through MAR (IVF/ICI, Intrauterine Insemination, and ovulation induction). We conducted a sensitivity analysis in which we considered only those children for whom information is available at all waves, and the results were virtually identical.

3.1. Measures

3.1.1. Psychosocial health

Our dependent variable is a measure of psychosocial problems derived from the Strengths and Difficulties Questionnaire (SDQ), which was first completed by the main interviewee when each child was around three years old (second wave) and at all subsequent waves. The SDQ – a standardised measure used to screen for psychiatric disorders (Goodman, 2001) – consists of 25 questions about five domains of behaviour: conduct problems, hyperactivity, emotional symptoms, peer problems, and pro-social behaviour (the complete list of items included in the Strengths and Difficulties Questionnaire is provided in the appendix). The main respondent could mark each of the attributes as "not true", "somewhat true", or "certainly true"; these responses are coded as zero, one, and two, respectively. The sub-scores for conduct problems, hyperactivity, emotional symptoms, and peer problems are summed together to create the total difficulties score with values ranging from zero to 40, whereby a higher value indicates that the child has more psychosocial problems. In addition, the first two sub-scores are summed to generate the *externalising* sub-scale, and the last two sub-scores are summed to generate the *internalising* sub-scale. Internalising problem behaviours include anxiety, depression, and psychosomatic reactions; and externalising problem behaviours include the outward expression of negative emotions through aggressive and delinquent behaviours. The scores can be interpreted either continuously or categorically, and are divisible into "normal" (total SDQ scores between 0–13), "borderline" (total SDQ scores between 14–16), or "abnormal" (total SDQ scores between 17–40) categories (Vanore, Mazzucato, & Siegel, 2015). We first consider the total difficulties score and then the two sub-scales. All of the scores are used as continuous variables.

3.1.2. Medically Assisted Reproduction (MAR)

The term MAR children refers to children who were conceived through assisted reproductive technology (ART) or other kinds of fertility treatments. The term ART refers to fertility treatments in which the fertilisation happens outside of the woman's body, such as IVF and ICSI. Other non-ART treatments of infertility include intrauterine insemination (IUI) and ovulation induction. In our analytical sample (N = 12,989), 404 children were conceived with the help of MAR. Of these, 180 were conceived through ART (either IVF or ICSI), 29 through IUI, and 195 with the aid of ovarian stimulation drugs not followed by any further treatment. In the main analyses, the different kinds of treatment are included in the same category.

3.1.3. Child and parental characteristics

We control for the basic characteristics of the cohort member: namely, the child's sex, and whether s/he was the first born. We also include adjustments for a number of confounders and mediators that could prove relevant for understanding the association between the use of MAR and children's psychosocial health. As confounders we control for the demographic and socio-economic characteristics of the parents: namely, the mother's age at birth (continuous), the mother's educational level (degree or equivalent, A-levels - which is the equivalent of high-school degree for those who plan to continue studying at university - or less than A-levels), the parents' marital status (married, cohabiting, single), the household's income quintile (obtained using the modified OECD scale, which adjusts family income by household size), and whether the pregnancy was planned (binary). As mediators, we control for whether the child had a low birth weight (defined as weighting less than 2.5 kg) and was delivered in a multiple birth. All of the confounders and the mediators are measured at the first wave, which took place when the cohort child was around nine months old; in 99% of the cases the cohort child's mother was the main respondent. In

additional analyses (available upon request), we tested whether the effect of MAR varied by sex of the child and birth order: none of these interaction effects was statistically significant and were therefore not included in the analyses. We also included an adjustment for parenting style and behaviours (a measure of maternal warmth – constructed using questions on the importance of cuddling, talking to and stimulating the child - and whether the parents read to the child every day), as these factors might mediate the association between the use of MAR and children's psychosocial health. However, the adjustment did not change the association. This finding is in line with the results of previous studies showing that the parenting styles and behaviours observed among MAR parents are fully explained by their socio-economic background (Barbuscia, 2017). Thus, the controls for maternal warmth and reading to the child were not included in the final analyses.

3.2. Methods

First, we display the descriptive characteristics of the sample divided by whether the cohort member was conceived through MAR or naturally, using two-tailed t-tests to assess whether the characteristics of the parents and of the children in the two sub-samples differ significantly. To investigate how psychosocial health develops as the children grow older, we also show the average levels of SDQ total difficulties scores and the internalising and externalising scores at each wave for the NC and the MAR children.

To explore whether and, if so, to what extent the psychosocial health of the MAR children differs from that of the NC children, we implemented multilevel random intercept models that allow us to account for the longitudinal structure of the data. An advantage of using such models is that the observations can be unbalanced; which in our case means that we are able to include children for whom observations are missing for one or more waves. We added interaction terms for the use of MAR and the age of the child to study whether and, if so, how the use of fertility treatment is associated with the psychosocial health of the child over time. Model 1 shows the unadjusted association between the use of MAR and the child's psychosocial health, controlling for baseline child characteristics. We then added covariates to test whether and, if so, how the unadjusted association changes after we take the confounders (parents' demographic and socio-economic characteristics) and the mediators (whether the child was first born and had a low birth weight) into account. Our main interest lies in estimating the coefficient of MAR and of its interaction with age, which shows how the association between the use of MAR and children's psychosocial outcomes (measured through SDQ) develop as the children grow up.

A general specification of the model is thus:

$$SDQ_{ij} = \alpha + \beta_{1*}MAR_i + \beta_{2*}MAR_{i*}age_j + \beta_{3*}X_i + \gamma_*age_j + \gamma^2 age_j + \delta_i + \varepsilon_{ij}$$

Where SDQ_{ij} is the difficulties score for child i at time j; MAR_i is an indicator that is one if the child was conceived through MAR and is zero otherwise; and age_j is the age of the child i at wave j (when the SDQ scores are measured). Since there is considerable variation in the ages of the children (on average, there is a 20-month difference between the youngest and the oldest child), age_j is the exact age of the child in years at each interview. X_i is a time-constant vector of individual child and parental characteristics. The coefficient β1 is the association between MAR and the baseline level of psychosocial health (at age three); β_2 indicates how the association develops with the age of the child; γ and γ^2 are the linear and quadratic effects of age; δ_i is the random intercept; and ε_{ij} is a time-specific error term. All of the predictors included in the models are time-invariant, and the characteristics were measured at the first wave. We first apply the model to the total SDQ difficulties score, and then separately to the internalising and externalising sub-scores. In all models, standard-errors are clustered to allow for non-independence of multiple births within families.

Table 1
Descriptive characteristics of the sample, by mode of conception.

	Natural conception % (se)	MAR % (se)
Multiple birth	2.0 (0.001)	22.7 (0.020)[***]
LBW	2.5 (0.001)	9.9 (0.013)[***]
First born	41.8 (0.00)	66.5 (0.02)[***]
Pregnancy was planned	55.4 (0.004)	100 (0)[***]
Mother's age (mean)	29.4 (0.052)	33.2 (0.215)[***]
Married parents	58.8 (0.003)	84.5 (0.018)[***]
Cohabiting parents	24.6 (0.003)	11.8 (0.014)
Single mother	15.7 (0.003)	3.7 (0.015)[***]
Degree or equivalent	17.8 (0.002)	27.5 (0.018)[***]
A levels	19.67(0.002)	23.7 (0.017)[*]
GCSE levels or lower	62.5 (0.004)	48.8 (0.021)
UK 1st quintile	21.1 (0.003)	4.4 (0.010)[***]
UK 2nd quintile	21.4 (0.003)	16.2 (0.015)[**]
UK 3rd quintile	19.9 (0.002)	20.6 (0.017)
UK 4th quintile	19.8 (0.002)	25.4 (0.018)[***]
UK 5th quintile	17.8 (0.002)	33.4 (0.20)[***]
Sample	12,585	404

Notes: Standard errors in parentheses. Stars indicate that the difference is significant according to two-tailed t-tests.
* p < 0.05.
** p < 0.01.
*** p < 0.001.

4. Results

Table 1 shows that there are consistent and statistically significant (according to two-tailed t-tests) differences between the MAR and the NC children. Almost 23% of the MAR children but only 2% of the NC children were delivered in a multiple birth. The MAR children also have higher probabilities of being the first born and of having low birth weight. Moreover, the demographic and socio-economic backgrounds of the MAR parents differ markedly from those of the parents who conceived naturally. Compared to the mothers who conceived naturally, the mothers who used MAR are, on average, older at the time of birth (33 years compared to 29 years), more likely to be married, more likely to be highly educated (25% have a degree or equivalent, compared to 15%). The MAR children are also more likely to come from a family with a high household income. The descriptive results show that the MAR children differ from their naturally conceived peers in terms of their outcomes at birth and their socio-economic backgrounds.

Table 2 shows that the children tended to have higher total difficulties scores when they were three years old (first observation) compared to older ages. As the scores measure the level of psychosocial problems, this finding suggests that the children had relatively low levels of psychosocial health at age three. The scores further decreased when the children were five years old, and then increased again when they were 14 years old. MAR children had lower scores than the NC children at each wave – differences were small but statistically significant. The two measures of internalising and externalising sub-scores developed differently with the ages of the children: the externalising sub-scores decreased consistently over time, while the internalising sub-scores decreased between the ages of three and five and then increased again after the age of seven. It appears that the differences between the MAR and the NC children were consistently higher on the externalising than on the internalising sub-scale.

4.1. Multilevel analysis

Table 3 displays the results of the multilevel random intercept models we used to study the association between the use of fertility treatments and the psychosocial development of the children in our sample from age three to age 14 (Table A1 in the appendix shows the full set of covariates). Results are shown for the total difficulties score

Table 2

Mean SDQ total difficulties scores, internalising and externalising scores, by age and by mode of conception.

SDQ scores	Natural conception mean (SD)	N	MAR mean (SD)	N
Total difficulties				
Age 3	9.66 (0.045)	12, 585	9.16* (0.236)	404
Age 5	7.39 (0.042)	11,400	6.78** (0.218)	369
Age 7	7.64 (0.048)	10,433	6.96** (0.246)	338
Age 11	7.31 (0.053)	10,574	5.85*** (0.296)	353
Age 14	7.70 (0.058)	9,710	6.85** (0.2844)	323
Internalising				
Age 3	2.91 (0.021)	12, 585	2.95 (0.125)	404
Age 5	2.52 (0.021)	11,400	2.41 (0.118)	369
Age 7	2.76 (0.025)	10,433	2.70 (0.141)	338
Age 11	3.01 (0.029)	10,574	2.43** (0.175)	353
Age 14	3.52 (0.033)	9,710	3.32 (0.171)	323
Externalising				
Age 3	6.75 (0.032)	12, 585	6.21** (0.166)	404
Age 5	4.87 (0.029)	11,400	4.36** (0.151)	369
Age 7	4.87 (0.032)	10,433	4.26** (0.15)	338
Age 11	4.24 (0.033)	10,574	3.07*** (0.185)	353
Age 14	4.11 (0.034)	9,710	3.40*** (0.170)	323

Notes: Stars indicate that the difference is significant according to two-tailed *t*-tests.

  * $p < 0.05$.

 ** $p < 0.01$.

*** $p < 0.001$.

and the externalising and internalising sub-scores. The coefficients of *MAR* represent the association between the use of treatments and the baseline level of psychosocial health (at age three), and the coefficients of the interaction between *MAR* and *age (MAR*age)* indicates whether the psychosocial development of MAR children differed from that of NC children over time.

The negative coefficient of *age* and the positive coefficient of *age2* suggest that, on average, the children's SDQ scores decreased (indicating better psychosocial health) as they grew older; but this trend is not linear. In Model 1 for the total difficulties score, the negative coefficient of *MAR* indicates that the MAR children had lower scores than the NC children at age three, although this association is not statistically significant. The significant and negative coefficient of the interaction term between MAR and age suggests that the psychosocial health gap between the MAR and the NC children widened with age; i.e., that the MAR children fared better than the NC children over time.

However, after we include controls for parental characteristics (i.e., the confounders, model 2), the association between the *MAR* and the total difficulties score at baseline becomes positive and statistically significant, which means that there is a negative association between MAR and children's psychosocial health. Adjusting for whether the child was born with a low birth weight or from a multiple pregnancy (i.e., the mediators, model 3) produced very small changes in the coefficients of interest. The results therefore suggest that the parental background of the MAR children accounts for the higher levels of psychosocial health observed among them; and that, net of parental background, MAR children exhibit worse psychosocial health. The size of the difference is small (the coefficient of MAR is around 0.9 on the score at age three, which is approximately 0.2 of a standard deviation). However, to give an example of how it compares to the association with other factors, the effect size is similar, but with the opposite sign, to the one of having a mother with at least A-level education compared to having a mother with less than A-level education. The coefficient of the interaction term remains negative and statistically significant in the adjusted model, which suggests that the differences in psychosocial outcomes attributable to fertility treatments attenuated; i.e., that the adjusted differences between the two groups converged as the children grew up. The results for the externalising and the internalising

Table 3

SDQ total difficulties, externalising and internalising scores and MAR, waves 2–6, estimates from multilevel random intercept models.

VARIABLES	Model 1: unadjusted	Model 2: Adjusted for confounders	Model 3: Fully adjusted
Total difficulties score			
MAR	− 0.363	0.735**	0.873***
	(0.285)	(0.277)	(0.281)
MAR*age	− 0.059**	− 0.061**	− 0.0622**
	(0.023)	(0.023)	(0.023)
Age	− 1.045**	− 1.038***	− 1.038***
	(0.023)	(0.023)	(0.023)
Age2	0.054***	0.054***	0.0540***
	(0.009)	(0.001)	(0.001)
Internalising sub-score			
MAR	0.086	0.408**	0.490***
	(0.152)	(0.151)	(0.153)
MAR*age	− 0.037**	− 0.038**	− 0.039**
	(0.014)	(0.014)	(0.014)
Age	− 0.202***	− 0.200***	− 0.200***
	(0.013)	(0.013)	(0.013)
Age2	0.016***	0.015***	0.015***
	(0.000)	(0.000)	(0.000)
Externalising sub-score			
MAR	− 0.397*	0.380*	0.484**
	(0.187)	(0.183)	(0.185)
MAR*age	− 0.035*	− 0.037**	− 0.037**
	(0.015)	(0.015)	(0.015)
Age	− 842***	− 0.839***	− 0.839***
	(0.015)	(0.015)	(0.015)
Age2	0.037***	0.037***	0.037***
	(0.000)	(0.000)	(0.000)
Observations	56,068	56,068	56,068
Number of groups	12,989	12,989	12,989

Standard errors in parentheses.

Wave takes a value from two to six. Model 1 controls for the sex of the child. Model 2 controls for parity, maternal age, education and marital status, HH income, and whether the pregnancy was planned. Model 3, in addition to the variables included in Model 2, controls for low birth weight, and whether the child was delivered in a multiple birth.

  * $p < 0.1$.

 ** $p < 0.05$.

*** $p < 0.01$.

problems sub-scales are similar. The only exception is that the coefficient on the intercept is negative and significant in the unadjusted model for the externalising sub-score, while it is positive and not statistically significant in the model for the internalising sub-score. The coefficients of the interactions are negative and statistically significant in all of the models, which suggests that the levels of internalising and externalising problems for MAR children followed a similar trend (and a pattern similar to the one observed for the total difficulties score).

The predicted total difficulties scores by the children's ages resulting from Models 1 and 3 are shown in Figs. 1 and 2. On average, within each sweep of data collection, there is a 20-months difference between the youngest and the oldest child. Therefore, in the Figures we plot the results using the exact ages of the children at each interview (and not the average age of children at each sweep i.e. ages 3, 5, 7, 11 and 14). As the coefficients of the multilevel random intercept models suggest, the MAR children had lower total difficulties scores at all ages before parental characteristics or outcomes at birth were taken into account. The difference was statistically significant and increased as the children grew older. However, these trends changed when we controlled for the potential confounding and mediating factors. The predicted values resulting from the adjusted model showed a converging trend over time.
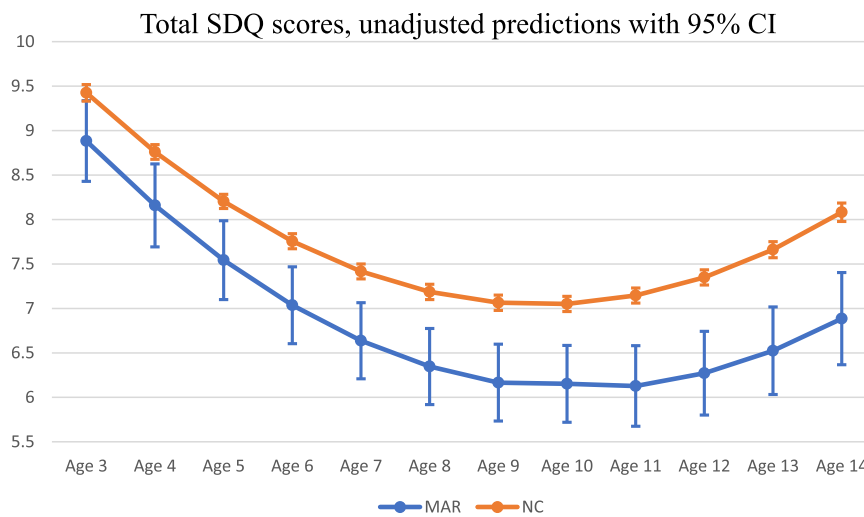
Fig. 1. Predicted psychosocial health among MAR and NC children, ages 3–14. Estimates from unadjusted random intercept models (Model 1).

While the MAR children had significantly higher scores up to the age of six, the scores converged thereafter: the differences in the SDQ scores of the MAR and the NC children become statistically and substantively negligible by end of follow-up at age 14.

These results suggest that, on average, the MAR children had significantly lower total, internalising, and externalising SDQ scores than the NC children from age four onwards. However, this difference could be explained by the advantaged socio-economic characteristics of the MAR parents. Once such characteristics were controlled for, the use of fertility treatment was found to be associated with lower levels of psychosocial health at ages three and five. These results are confirmed by additional analyses performed on the sample restricted to families with income in the 4th or 5th quintile of the UK household income distribution (Appendix Table A3). The use of fertility treatment was also shown to be associated with decreasing SDQ scores over time in both the adjusted and the unadjusted models.

5. Conclusion and discussion

The scarcity of studies looking at the longer-term development of MAR children represents a major gap in the existing research on the outcomes of MAR. Filling this gap in our knowledge is especially relevant in the current contest, as the numbers of individuals and couples

using fertility treatments to help them conceive – and thus the numbers of children born through MAR – are increasing (Calhaz-Jorge et al., 2016). While the existing evidence suggests that MAR children are not, on average, at increased risk of having poor psychosocial health during childhood, the evidence on the longer-term outcomes of MAR children is more limited. It is important that we investigate the health of these children as they enter adolescence, as psychosocial problems and depressive symptoms often emerge in children around the age of 14 (Patalay & Fitzsimons, 2017). Furthermore, several scholars have suggested that an MAR child's development and relationship with her/his parents may enter a particularly difficult phase during adolescence, as the parents might wait until the child reaches early adolescence to reveal to their daughter or son how s/he was conceived (Colpin & Soenen, 2002; Golombok et al., 1996; Ilioi & Golombok, 2015).

There are two main reasons why it is important to investigate whether the use of MAR is associated with lower levels of psychosocial health in children. First, there is evidence that children with a higher SDQ score exhibit higher levels of psychopathology at both the time of the test and later in life (Goodman, 2001). Second, the non-cognitive abilities individuals develop in childhood have been shown to be important predictors of a number of outcomes at older ages, including their educational achievement, earnings, and employment status (among other outcomes, see Heckman, Stixrud, & Urzua, 2006).
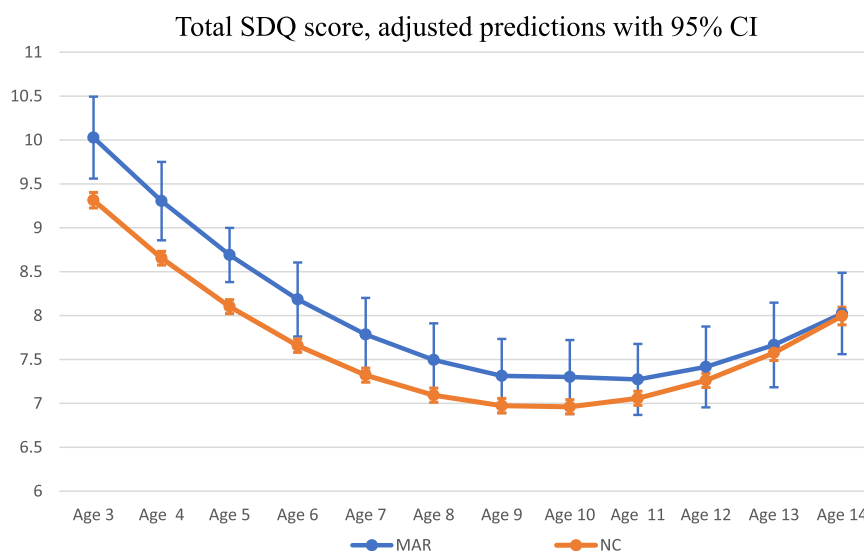

Fig. 2. Predicted psychosocial health among MAR and NC children, ages 3–14. Estimated from fully adjusted random intercept models (Model 2).

Using data from six waves of the UK MCS, we compared the psychosocial health (as measured by the SDQ total difficulties score and the internalising and externalising sub-scales) of children conceived through MAR to that of their naturally conceived peers up to the age of 14. Taking a longitudinal perspective is important because cognitive and non-cognitive skills develop over time starting in early childhood. However, the existing literature on this topic has focused on children's outcomes at a single point in time, and often at very young ages. The availability of detailed information on the demographic and socio-economic characteristics of parents allowed us to control for a number of important potential confounders. Controlling for parental characteristics is especially relevant in the UK context, as MAR treatments are very expensive (the average price for one complete treatments cycle is around 3,000 £ for IVF and 4,000 £ for ICSI) and often not publicly funded. The individuals who have access to fertility treatments tend thus to be a selected subgroup of the population, which is itself likely to be an important determinant of children's outcomes.

While concerns are often raised about the development of children conceived through fertility treatments, we found that MAR children do not suffer from worse psychosocial health than NC children. Our results show that, before parental characteristics are taken into account, the association between the use of the MAR and the SDQ total difficulties, internalising, and externalising scores at age three (when they are first observed) is negative, but not statistically significant. The association becomes negative and significant at older ages, which indicates that MAR children exhibit higher levels of psychosocial health than NC children over time. However, after controlling for parental socio-economic characteristics, the association is inverted, and the use of fertility treatment is found to be associated with worse psychosocial health in children up to age five. Yet the differences observed are small, and are thus unlikely to result in a higher incidence of psychiatric disorders. Moreover, the findings indicate that the use of fertility treatments is associated with increasing levels of psychosocial health over time, and that the levels in the two groups of children fully converge at age 14 (the last observation). Thus,we found no evidence that MAR children are at higher risk of developing psychosocial problems when they reach early adolescence, as has been hypothesised in the literature (Colpin & Soenen, 2002; Hart & Norman, 2013b; Wagenaar et al., 2008). While we were not able to directly test the potential role of the disclosure of the way of conception on children's psychosocial health in early adolescence (as this information is not collected in the MCS), our results could suggest that disclosure (or lack thereof) does not result in higher psychosocial problems.

Mothers who conceived with the aid of fertility treatments represent a selected population, as they are, on average, older and have more advantaged socio-economic backgrounds than other parents. These selected characteristics seem to explain the relatively high levels of psychosocial health observed among MAR children.

A possible explanation for our finding that there is a negative association between the use of MAR and psychosocial health at age three is that a mother who conceived through MAR may be more anxious about her child's health in the first years of life. This might be due, for example, to worries about health issues related to sub-fertility and the MAR treatments, and concerns about the health of the newborn child, who might be perceived as being especially fragile in the first years of life. That mothers who conceived through MAR might experience higher levels of overprotection towards their toddler has been suggested by a number of studies (Colpin & Soenen, 2002; Gibson et al., 2000; Hahn & Di Pietro, 2001). While these studies did not directly test whether this may vary (decrease) over time as the children get older, our finding that the adjusted differences in the psychosocial health of MAR and NC children converge over time could be explained by a tendency for these worries to become less salient as the time since conception passes. However, the MCS does not enable us to test these hypotheses. Alternatively, the results could reflect the finding that – after accounting for background characteristics – children conceived through MAR children have worse psychosocial health in childhood. A plausible explanation for this finding is that these effects are related to the worse birth outcomes of MAR children; however, an adjustment for LBW did not change the results.

The study suffers from some limitations. First, the sample size of MAR children is small which prevents us from running separate analyses by MAR treatments. Nonetheless, additional analyses (available upon request) show that the patterns are qualitatively similar when we disaggregate the MAR group by treatment type. Second, we rely on mother-reported measures, which might reflect factors such as higher levels of concerns for the children's health in their first years of life. Third, with the information available in the MCS data, we were unable to directly test the mechanisms explaining the higher levels of psychosocial problems among MAR children in the adjusted models in the first years of life. However, these limitations do not overshadow the key strengths of this study. First, information about the level of psychosocial problems at different points in time permitted us to explore the development of children over time and to show – contrary to what has been hypothesized by previous studies - that MAR children are not at higher risk of developing psychosocial problems when they reach early adolescence. Second, the inclusion of rich information about parental background enabled us to reach a deeper understanding of the potential factors affecting MAR children's development and whether it differs from that of naturally conceived children.

Our finding that, on average, MAR children are not at higher risk than NC children of developing poor psychosocial health during and after childhood is important, especially given the concerns expressed about the outcomes of MAR, and the lack of consistent evidence on this relationship over the longer term. Furthermore, the results of this research are especially relevant at a time when an increasing number of MAR children are entering adolescence or adulthood.

Conflict of interest

None declared.

Appendix

Strengths and Difficulties Questionnaire (SDQ)

For each item, respondents are asked to indicate whether it is Not True, Somewhat True or Certainly True, on the basis of the child's behavior in the previous six months

- Considerate of other people' s feelings
- Restless, overactive, cannot stay still for long
- Often complain of headaches, stomach-aches or sickness
- Shares readily with other children (treats, toys, pencils etc.)
- Often has temper tantrums of hot tempers
- Rather solitary, tends to play alone
- Generally obedient, usually does what adults request

- Many worries, often seems worried
- Helpful if someone is hurt, upset or ill
- Constantly fidgeting or squirming
- Has at least one good friend
- Often fights with other children or bullies them
- Often unhappy, down-hearted or tearful
- Generally liked by other children
- Easily distracted, concentration wanders
- Nervous or clingy in new situations, easily loses confidence
- kind to younger children
- often lies or cheats
- picked on or bullied by other children
- often volunteers to help others (parents, teachers, other children)
- thinks things before acting
- steals from home, school or elsewhere
- gets on better with adults than with other children
- many fears, easily scared
- sees tasks through the end, good attention span

.

See Tables A1–A3 here.

Appendix Table A1
MAR and SDQ total difficulties scores. Estimates from random intercept models.

VARIABLES	Total SDQ		
	Model 1	*Model 2*	*Model 3*
MAR	−0.363	0.735***	0.873***
	(0.285)	(0.277)	(0.281)
MAR*age	−0.059**	−0.061***	−0.062**
	(0.023)	(0.023)	(0.023)
Age	−1.045***	−1.038***	−1.038***
	(0.023)	(0.023)	(0.023)
Age2	0.054***	0.054***	0.054***
	(0.001)	(0.001)	(0.001)
Female	−1.094***	−1.097***	−1.099***
	(0.076)	(0.071)	(0.071)
First born		0.143*	0.145*
		(0.080)	(0.080)
Maternal age		0.143*	0.145*
		(0.080)	(0.080)
Marital status (ref: single)			
Married		−0.705***	−0.705***
		(0.126)	(0.126)
Cohabiting		−0.173	−0.175
		(0.126)	(0.126)
Planned pregnancy		−0.334***	−0.336***
		(0.081)	(0.081)
Educational level (ref: less than A levels)			
Degree or higher		−1.420***	−1.412***
		(0.111)	(0.110)
A levels		−0.920***	−0.914***
		(0.096)	(0.096)
Household income (ref: UK income 5th quintile)			
1st quintile		2.220***	2.231***
		(0.150)	(0.150)
2nd quintile		1.366***	1.388***
		(0.132)	(0.133)
3rd quintile		0.585***	0.603***
		(0.125)	(0.125)
4th quintile		0.080	0.100
		(0.118)	(0.118)
Multiple birth			−1.032***
			(0.240)
LBW			1.038***
			(0.229)

Appendix Table A1 (*continued*)

VARIABLES	Total SDQ		
	Model 1	*Model 2*	*Model 3*
Constant	12.61***	14.36***	14.31***
	(0.097)	(0.306)	(0.306)
Sd(cons)	3.977	3.632	3.627
	(0.030)	(0.028)	(0.028)
Sd(residual)	3.674	3.683	3.682
	(0.012)	(0.012)	(0.012)
Observations	56,068	56,068	56,068
Number of groups	12,585	12,585	12,585

Standard errors in parentheses.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

Appendix Table A2
MAR and SDQ internalizing and externalizing difficulties scores. Estimates from random intercept model.

	Internalizing problems			Externalizing problems		
	Model 1	*Model 2*	*Model 3*	*Model 1*	*Model 2*	*Model 3*
MAR	0.086	0.408***	0.490***	−0.397**	0.380**	0.484***
	(0.153)	(0.151)	(0.153)	(0.188)	(0.183)	(0.185)
MAR*age	−0.037***	−0.038***	−0.039***	−0.035**	−0.037**	−0.037**
	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)
Age	−0.203***	−0.200***	−0.200***	−0.843***	−0.839***	−0.839***
	(0.013)	(0.013)	(0.013)	(0.015)	(0.015)	(0.015)
Age2	0.016***	0.0160***	0.0160***	0.037***	0.037***	0.037***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female	−0.041	−0.044	−0.043	−1.062***	−1.059***	−1.057***
	(0.037)	(0.036)	(0.036)	(0.050)	(0.047)	(0.047)
First born		0.287***	0.289***		−0.144***	−0.141***
		(0.040)	(0.040)		(0.053)	(0.053)
Maternal age		−0.004	−0.0036		−0.051***	−0.050***
		(0.003)	(0.003)		(0.004)	(0.004)
Marital status (ref: single)						
Married		−0.139**	−0.137**		−0.569***	−0.566***
		(0.063)	(0.063)		(0.083)	(0.083)
Cohabiting		−0.090	−0.089		−0.078	−0.077
		(0.063)	(0.063)		(0.083)	(0.083)
Planned pregnancy		−0.155***	−0.156***		−0.183***	−0.184***
		(0.041)	(0.041)		(0.051)	(0.054)
Educational level (ref: less than A levels)						
Degree		−0.397***	−0.393***		−1.018***	−1.014***
		(0.055)	(0.055)		(0.073)	(0.073)
A levels		−0.337***	−0.333***		−0.579***	−0.575***
		(0.048)	(0.048)		(0.064)	(0.064)
Household income (ref: UK income 5th quintile)						
1st quintile		1.187***	1.196***		1.027***	1.039***
		(0.075)	(0.075)		(0.098)	(0.099)
2nd quintile		0.805***	0.818***		0.557***	0.575***
		(0.066)	(0.066)		(0.087)	(0.087)
3rd quintile		0.390***	0.401***		0.192**	0.206**
		(0.063)	(0.063)		(0.082)	(0.082)
4th quintile		0.0627	0.0738		0.0183	0.0325
		(0.059)	(0.059)		(0.078)	(0.078)
Multiple birth			−0.554***			−0.700***
			(0.121)			(0.158)
LBW			0.463***			0.554***
			(0.116)			(0.152)
Constant	3.335***	3.148***	3.119***	9.295***	11.23***	11.19***

Appendix Table A2 (*continued*)

	Internalizing problems			Externalizing problems		
	Model 1	*Model 2*	*Model 3*	*Model 1*	*Model 2*	*Model 3*
	(0.0553)	(0.156)	(0.157)	(0.0643)	(0.202)	(0.202)
Sd (cons)	1.847	1.738	1.735	2.608	2.401	2.398
	(0.015)	(0.015)	(0.015)	(0.019)	(0.018)	(0.018)
Sd(residuals)	2.201	2.199	2.199	2.429	2.428	2.428
	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
Observations	56,166	56,068	56,068	56,166	56,068	56,068
Number of groups	12,989	12,965	12,965	12,989	12,965	12,965

Standard errors in parentheses.

*p < 0.1.

** p < 0.05.

*** p < 0.01.

Appendix Table A3

MAR and SDQ total difficulties scores. Estimates from random intercept model, sample restricted to households with income in the 4th and 5th UK income distribution quintiles.

VARIABLES	Model (1) Tot SDQ	Model (2) Tot SDQ	Model (3) Tot SDQ
MAR	0.427	0.504	0.590*
	(0.314)	(0.315)	(0.321)
MAR * age	−0.0522**	−0.0524**	−0.0527**
	(0.0266)	(0.0265)	(0.0265)
Age	−0.914***	−0.913***	−0.913***
	(0.0313)	(0.0313)	(0.0313)
Age 2	0.0471***	0.0471***	0.0471***
	(0.00183)	(0.00183)	(0.00183)
Female	−1.012***	−0.959***	−0.960***
	(0.103)	(0.100)	(0.100)
First born		0.454***	0.454***
		(0.107)	(0.107)
Maternal age		−0.0457***	−0.0452***
		(0.0121)	(0.0121)
Marital status (ref: single)			
Married		−1.349***	−1.326***
		(0.326)	(0.326)
Cohabiting		−1.004***	−0.993***
		(0.339)	(0.339)
Planned pregnancy		−0.128	−0.137
		(0.120)	(0.120)
Educational level (ref: less than A levels)			
Degree		−1.282***	−1.277***
		(0.121)	(0.121)
A levels		−0.547***	−0.539***
		(0.128)	(0.128)
Multiple birth			−0.745**
			(0.361)
LBW			0.789**
			(0.358)
Constant	10.58***	13.76***	13.72***
	(0.133)	(0.526)	(0.526)
Sd(cons)	2.987	2.522	2.527
	(0.020)	(0.018)	(0.018)
Sd(residuals)	2.464	2.582	2.571
	(0.011)	(0.011)	(0.011)
Observations	22,292	22,292	22,292
Number of groups	4,962	4,962	4,962

Standard errors in parentheses.

* p < 0.1.

** p < 0.05.

*** p < 0.01.

References

Barbuscia, A., & Mills, M. (2017). Cognitive development in children up to age 11 years born after ART- a longitudinal cohort study. *Human Reproduction, 32*(7), 1482–1488.

Barbuscia, Anna (2017). *Outcomes of fertility treatments for children and families.* Oxford University (Evidence from the UK and the US (PhD thesis)).

Bay, B., Mortensen, E. L., Hvidtjorn, D., & Schioler, U. (2013). Fertility treatment and risk of childhood and adolescent mental disorders: Register based cohort study. *British Medical Journal, 347,* 3978.

Bernstein, Judith (1990). Parenting after Infertility. *The Journal of Perinatal and Neonatal Nursing, 4*(2), 11–23.

Boardman, J. D., Powers, D. A., & Hummer, R. A. (2002). Low birth weight, social factors, and developmental outcomes among children in the United States. *Demography, 39*(2), 353–368.

Calhaz-Jorge, C., et al. (2016). Assisted reproductive technology in Europe, 2012: Results generated from European registers by ESHRE. *Human Reproduction, 31*(8), 1638–1652.

Carson, C., Kelly, Y., Kurinczuk, J. J., Sacker, A., Redshaw, M., & Quigley, M. A. (2011). Effect of pregnancy planning and fertility treatment on cognitive outcomes in children at ages 3 and 5: Longitudinal cohort study. *British Medical Journal, 343,* 44–73.

Carson, C., Redshaw, M., Sacker, A., Kelly, Y., Kurinczuk, J. J., & Quigley, M. A. (2013). Effects of pregnancy planning, fertility, and assisted reproductive treatment on child behavioral problems at 5 and 7 years: Evidence from the Millennium Cohort Study. *Fertility and Sterility, 99*(2), 456–463.

Chambers, G. M., Hoang, V. P., Sullivan, E. A., Chapman, M. G., Ishihara, O., Zegers-Hochschild, F., & Adamson, G. D. (2014). The impact of consumer affordability on access to assisted reproductive technologies and embryo transfer practices: An international analysis. *Fertility and Sterility, 101*(1), 191–198.

Chandra, Anjani, Copen, C. E., & Stephen, E. H. (2014). *Infertility service use in the United States: Data from the national survey of family growth, 1982–2010.* 1–21 (National health statistics reports).

Colpin, H., & Bossaert, G. (2008). Adolescents conceived vy IVF: Parenting and psychosocial adjustement. *Human Reproduction, 13*(12), 2724–2730.

Colpin, H., & Soenen, S. (2002). Parenting and psychosocial development of IVF children: A follow-up study. *Human Reproduction, 17*(4), 1116–1123.

Colpin, Hilde, (2002). "Parenting and Psychosocial Development of IVF Children: Review of the Research Literature", 22(4): 644–73.

Conti, G., & Heckman, J. J. (2014). Economics of Child Well-Being. In A. Ben-Arieh, F. Casas, I. Frønes, & J. Korbin (Eds.). *Handbook of Child Well-Being.* Dordrecht: Springer.

Cunha, F., & Heckman, J. (2007). The technology of skill formation author(s): source: T. *American Economic Review, 97,* 31–47.

Cunha, F., Heckman, J., & Schennach, S. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica, 78*(3), 1–47.

Ermisch, John (2008). Origins of social immobility and inequality: parenting and early child development. *National Institute Economic Review, 205*(1), 62–71.

Feinstein, Leon (2002). Inequality in the early cognitive development of British children in the 1970 cohort. *Economica, 70*(277), 73–97.

Gibson, F. L., Ungerer, J. A., Leslie, G. I., Leslie, Saunders, D. M., & Tennant, C. C. (1998). Development, behavior and temperament: A prospective study of infants conceived through in-vitro fertilization. *Human Reproduction, 13*(6), 1727–1732.

Goodman, R. (2001). Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(11), 1337–1345.

Golombok, S., Blake, L., Casey, P., Roman, G., & Jadva, V. (2013). Children born through reproductive donation: A longitudinal study of psychological adjustment. *The Journal of Child Psychology and Psychiatry, 54*(6), 653–660.

Golombok, S., Brewaeys, A., Cook, R., Giavazzi, M. T., Guerra, D., Mantovani, A., van Hall, E., Crosignani, P. G., & Dexeus, S. (1996). The European study of assisted reproduction families: Family functioning and child development. *Human Reproduction, 11*(10), 2324–2331.

Golombok, S., Brewaeys, A., Giavazzi, M. T., Guerra, D., MacCallum, F., & Rust, J. (2002). The European study of assisted reproduction families: The transition to adolescence. *Human Reproduction, 17*(3), 830–840.

Golombok, S., Cook, R. a, Bish, & Murray, C. (1995). Families created by the new reproductive technologies: Quality of parenting and social and emotional development of the children. *Child development, 66*(2), 285–298.

Golombok, S., Ilioi, E., Blake, L., Roman, G., & Jadva, V. (2017). A longitudinal study of families formed through reproductive donation: Parent-adolescent relationships and adolescent adjustment at age 14. *Developmental Psychology, 53*(10), 1966–1977.

Golombok, S., MacCallum, F., & Goodman, E. (2001). The 'test-Tube' generation: Parent-Child relationships and the psychological well-being of in vitro fertilization children at adolescence. *Child development, 72*(2), 599–608.

Hahn, C.-S., & DiPietro, J. (2001). In vitro fertilization and the family: Quality of parenting, family functioning, and child psychosocial adjustment. *Developmental Psychology, 37,* 37–48.

Hart, R., & Norman, R. J. (2013a). The longer-term health outcomes for children born as a result of IVF treatment: Part I–General health outcomes. *Human Reproduction Update, 19*(3), 232–243.

Hart, R., & Norman, R. J. (2013b). The longer-term health outcomes for children born as a result of IVF treatment. Part II–Mental health and development outcomes. *Human Reproduction Update, 19*(3), 244–250.

Heckman, James J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900–1902.

Heckman, James J., Stixrud, J., & Urzua, s (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*(3), 411–482.

Ilioi, E. C., & Golombok, S. (2015). Psychological adjustment in adolescents conceived by assisted reproduction techniques: A systematic review. *Human Reproduction Update, 21*(1), 84–96.

Levy-Shiff, R., Vakil, E., Dimitrovsky, L., et al. (1998). Medical, cognitive, emotional and behavioural outcomes in shool-age children conceived by in vitro fertilization. *Journal of Clinical Child Psychology, 27*(3), 320–329.

Lewis, H., Hope, S., & Pearce, A. (2014). Socioeconomic inequalities in parent-reported and teacher-reported psychological well-being. *Archives of Disease in Childhood, 0,* 1–14.

Ludwig, A. K., Sutcliffe, A. G., Diedrich, K., & Ludwig, M. (2006). Post-natal health and development of children born after assisted reproduction: A systematic review of controlled studies. *European Journal of Obstetrics and Gynecology and Reproductive Biology, 127*(1), 3–25.

Owen, L., & Golombok, S. (2009). Families created by assisted reproduction: Parent-Child Relationships in late adolescence. *Journal of Adolescence, 32*(4), 835–848.

Patalay, P., & Fitzsimons, E. (2017). *Mental ill-health among children of the new century: Trends across childhood with a focus on age 14.* London: Centre for Longitudinal Studies.

Pinborg, A., Henningsen, A.-K., Malhau, S., & Loft, A. (2013). Congenital anomalies after assisted reproductive technology. *Fertility and Sterility, 99*(2), 327–332.

Pison, G., Monden, C., & Smits, J. (2015). Twinning rates in developed countries: Trends and explanations. *Population and Development Review, 41*(4), 629–649.

Präg, P., & Mills, M. C. (2015). Assisted reproductive technologies in Europe: Usage and Regulation in the contex of cross-border reproductive care. In M. Kreyenfeld (Ed.). *Childlessness in Europe: Context, cause, and consequences* (pp. 289–312). Springer Open.

van Balen, Frank (1998). Development of IVF children. *Developmental Review, 18*(1), 30–46.

Vanore, M., Mazzucato, V., & Siegel, M. (2015). 'Left behind' but not left alone: Parental migration & the psychosocial health of children in Moldova. *Social Science Medicine, 132,* 252–260.

Wagenaar, K., Ceelen, M., Van Weissenbruch, M. M., Knol, D. L., Delemarre-Van De Waal, H. A., & Huisman, J. (2008). School functioning in 8- to 18-year-old children born after in vitro fertilization. *European Journal of Pediatrics, 167*(11), 1289–1295.

Wagenaar, K., van Weissenbruch, M. M., Knol, D. L., Cohen-Kettenis, P. T., Delemarre-van de Waal, H. A., & Huisman, J. (2009). Behavior and socioemotional functioning in 9-18-year-old children born after in vitro fertilization. *Fertility and Sterility, 92*(6), 1907–1914.

Wagenaar, K., van Weissenbruch, M. M., van Leeuwen, F. E., Cohen-Kettenis, P. T., Delamarre-van de Waal, H. A., Schats, R., & Huisman, J. (2011). Self-reported behavioral and socioemotional functioning of 11- to 18-year-old adolescents conceived by in vitro fertilization. *Fertility and Sterility, 95*(2), 611–616.

Wilson, C. L., Fisher, J. R., Hammarberg, K., Amor, D. J., & Halliday, J. L. (2011). Looking downstream: A review of the literature on physical and psychosocial health outcomes in adolescents and young adults who were conceived by ART. *Human Reproduction, 26*(5), 1209–1219.