Opposite-sex and same-sex twin studies of physiological, cognitive and behavioral traits

Ahrenfeldt, Linda Juel; Christensen, Kaare; Segal, Nancy L; Hur, Yoon-Mi

Published in:
Neuroscience & Biobehavioral Reviews

DOI:
10.1016/j.neubiorev.2019.11.004

Publication date:
2020

Document version:
Accepted manuscript

Opposite-sex and same-sex twin studies of physiological, cognitive and behavioral traits

A special issue for Neuroscience and Biobehavioral Reviews

Linda Juel Ahrenfeldt[1,2], Kaare Christensen[1,2,3,4], Nancy L. Segal[5] and Yoon-Mi Hur[6]

1) Unit of Epidemiology, Biostatistics and Biodemography, Department of Public Health, University of Southern Denmark, 5000 Odense, Denmark
2) The Danish Twin Registry, Department of Public Health, University of Southern Denmark, 5000 Odense, Denmark
3) Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, 5000 Odense, Denmark
4) Department of Clinical Genetics, Odense University Hospital, 5000 Odense, Denmark
5) Department of Psychology, California State University, 800 N. State College Blvd., Fullerton, CA, 92834, USA
6) Research Institute for the Welfare Society, Mokpo National University, Jeonnam, South Korea

Address for correspondence: Yoon-Mi Hur, Research Institute for the Welfare Society, Mokpo National University, Jeonnam, South Korea. Email: ymhur@mokpo.ac.kr

Abstract

A scientific interest in opposite-sex (OS) twins comes from animal studies showing hormone transfer between fetuses *in utero*. A parallel effect in humans may occur, especially for OS females who may be exposed to androgens, in particular testosterone, from the male co-twin. Conversely, OS males may be exposed to lower levels of prenatal testosterone than do same-sex (SS) males. In this special issue, we reviewed published studies investigating potential differences between OS and SS twins in physiological, cognitive and behavioral traits focusing on the Twin Testosterone Transfer hypothesis. Sixty articles fulfilled the eligibility criteria including 23 studies published since the review by Tapp et al. (2011). In general, studies of cognition are conflicting, but it is the phenotype for which most support for the TTT hypothesis is found. Less consistent evidence has been found regarding physiological and behavioral traits. We hope that this special issue will stimulate a discussion about how an investigation of the TTT hypothesis should continue in future research.

***Keywords*:** Opposite-sex; Same-sex; Twins; Sex differences; Testosterone; Masculinization; Physiology, Cognition; Behavior

Highlights:

Sex hormones may be transferred between fetuses in the same pregnancy (the Twin Testosterone Transfer hypothesis).

The TTT hypothesis was evaluated using 60 published studies investigating potential differences between opposite-sex and same-sex twins.

Cognition is the phenotype for which most support for the TTT hypothesis is found.

Studies of physiology and behavior are less supportive of the TTT hypothesis.

Overall, a growing body of research is challenging the manifestation of the TTT hypothesis.

1 Twins

Twin pregnancies are characterized by simultaneous development of two fetuses sharing the uterus. There are two kinds of twins: monozygotic (MZ) and dizygotic (DZ) twins. MZ twins are derived from one fertilized egg, from which two separate embryos later emerge and are genetically identical at conception (Hall, 2003). However, there are different subtypes of MZ twins, possible due to the timing of the initial zygotic division (Loos et al., 1998): dichorionic diamniotic (DCDA), monochorionic diamniotic (MCDA), and monochorionic monoamniotic (MCMA) twins. DCDA twins, who have separate chorions (placenta) and amnions, form when cleavage takes place within 72 hours of fertilization (about 1/3 of MZ twins). MCDA twins, who share the same chorion (placenta) but have two amnions, form when cleavage occurs within 4 to 7 days of fertilization (about 2/3 of MZ twins). MCMA twins, who share the same chorion and amnion, occurs when the split takes place within 8 to 14 days of fertilization (about 1-2% of MZ twins). As MCMA and MCDA twins share the same chorion, they have an increased risk for twin-to-twin transfusion syndrome as well as for perinatal mortality and morbidity compared with DCDA twins, who have separate chorions and amniotic sacs (Hall, 2003; Lewi, 2010; Loos et al., 1998). As DZ twins emerge from two different egg cells and two different sperm cells, they are always DCDA twins. Like ordinary siblings, DZ twins share about 50% of the segregated genes identical by decent (Hall, 2003). Contrary to MZ twins, who are always of the same sex, DZ twins can be either same-sex (SS) or opposite-sex (OS). Thus, the OS twins are always DZ whereas the SS twin pairs, consisting of two boys or two girls, can be either MZ or DZ.

2 The Twin Testosterone Transfer hypothesis

2.1 Sex hormones

Gonadal hormones, particularly androgens, play an important role in early human development, influencing both physical and behavioral characteristics (Hines, 2011). Testosterone is the major androgenic hormone produced by the testes. The fetal testes begin to produce testosterone prenatally, but the ovaries do not (Wilson et al., 1981). As a consequence, male fetuses are exposed to higher levels of testosterone than are female fetuses (Hines, 2008). Testosterone production in males is highest from 8-24 weeks of gestation, with peak levels occurring between 12-18 gestational weeks (Abramovich, 1974; Nagamani et al., 1979; Warne et al., 1977), and these periods are times of rapid brain development. Gonadal hormones can induce a masculinizing influence, enhancing the development of male-typical traits, or a feminizing influence, which enhances the development of female-typical traits. Gonadal hormones can also have a demasculinizing influence (suppression of male-typical characteristics) or a defeminizing influence (suppression of female-typical characteristics) (Collaer and Hines, 1995). Prenatal testosterone exposure has been linked to the masculinization of a variety of traits, including behavioral changes such as childhood play behavior, sexual orientation and gender identity, as well as some cognitive, motor and personality characteristics that show sex differences (Hines, 2006, 2010). In general, estrogen does not promote female-typical development. This occurs in the absence or reduction of testicular hormones (Hines, 2011).

2.2 Animal studies

Human studies of prenatal hormone effects were initially motivated by experimental studies in non-human animals. The pioneering study by Phoenix et al. found that female guinea pigs that were exposed to testosterone prenatally showed masculinized behavior in adulthood (Phoenix, 2009). Since then, numerous studies in non-human mammals have demonstrated effects of testosterone on neurobehavioral sexual differentiation (Constantinescu and Hines, 2012). Moreover, studies in, for example, rats and mice have demonstrated that exposure to sex hormones is influenced by the intrauterine fetal position of the animal, e.g. the proximity to fetuses of the same or opposite sex. Females are considered to be more sensitive to intrauterine position effects than males (Ryan and Vandenbergh, 2002) and female rodents developing between male fetuses *in utero* present masculinized anatomical (e.g. increased anogenital distance), behavioral (e.g. more aggressive behavior, less attractive to males) and reproductive characteristics (e.g. less reproductive) in adulthood compared with females developing between female fetuses (Ryan and Vandenbergh, 2002). Irrespective of sex, a fetus located between two male fetuses has higher blood concentrations of testosterone and lower concentrations of estradiol than fetuses located between two females (Ryan and Vandenbergh, 2002; vom Saal, 1989). However, although the effects of intrauterine positions on both sexes are well documented in model species, recent literature emphasizes the importance of examining intrauterine position effects in species with different life histories (Fishman et al., 2019). A recent study investigating the intrauterine position effects in a feral animal model (Myocastor coypus), which is characterized by long gestation and precocious offspring, found the opposite, namely that in rodent model species females adjacent to males *in utero* did not show increased testosterone levels. To the contrary, they showed a reduction in testosterone immunoreactivity, while the testosterone levels of females not positioned next to a male did not differ from those of males (Fishman et al., 2019).

2.3 Congenital Adrenal Hyperplasia

Evidence that testosterone influences human neurobehavioral development comes largely from studies of humans who develop in atypical hormonal environments (Constantinescu and Hines, 2012). The best-studied clinical condition is Congenital Adrenal Hyperplasia (CAH) (Cohen-Bendahan et al., 2005b). Individuals with CAH produce high levels of androgens from early in gestation, due to an enzymatic defect caused by a single gene (Pang et al., 1980). Females with CAH are suggested to differ from unaffected females in a number of domains including activity interests, personality, cognitive abilities, handedness, and sexuality (Hines et al., 2003). However, the most convincing evidence comes from studies of childhood play with CAH females, showing increased male-typical toy, playmate, and activity preferences (Hines, 2011). Males with CAH are generally found to be similar to their unaffected brothers with regard to most aspects of behaviors, but there is some suggestion that they have lower spatial ability than control males (Hampson et al., 1998; Hines et al., 2003).

2.4 Maternal-fetal and feto-fetal route

An ideal study from which to obtain information about the prenatal environment would be a study that directly measures fetal hormones at many points in gestation, and then follows the children into childhood and beyond. However, because there is risk associated with the collection of serum from living fetuses, this is not feasible (Cohen-Bendahan et al., 2005b). Alternative measures have included the investigation of androgen concentrations in maternal serum during pregnancy (Hickey et al., 2009) and perinatal hormones obtained from umbilical cord blood at birth (Whitehouse et al., 2010). It has also been suggested (Cohen-Bendahan et al., 2005b) that studies of prenatal hormones from the amniotic sac during the second trimester of pregnancy provide the most accurate measure of fetal androgen exposure, but this procedure is only performed when there is a medical reason, and the participant sample may not be representative of the general population (Tapp et al., 2011).

OS twins are thought to provide another opportunity to test the effects of prenatal testosterone exposure (Miller, 1994). The Twin Testosterone Transfer (TTT) hypothesis assumes that human sex hormones are transferred between fetuses in the same pregnancy of twins. Transfer of testosterone may occur via the maternal circulation or more directly between fetuses (Miller, 1994). In the first route, testosterone is suggested to pass from fetus to fetus through maternal circulation. Empirical evidence for this pathway comes from animal studies showing that testosterone injected into a pregnant mother increased circulating testosterone concentrations in the fetus and exerted a masculinizing influence on offspring behavior (Miller, 1994; Phoenix et al., 1959). In humans, however, the suggestion of a maternal route is not supported because hormone levels in maternal blood and amniotic fluid do not correlate (Nagamani et al., 1979). Moreover, other human studies have found that the sex of a fetus cannot be predicted from serum androgen concentrations (Glass and Klein, 1981; Hines et al., 2002; van de Beek et al., 2004) suggesting that maternal-fetal hormone transfer may be unidirectional, with only passage of hormones from the mother to the fetus (Tapp et al., 2011). The other potential route goes directly from fetus to fetus (diffusion across fetal membranes). Amniotic fluid can permeate the fetal skin and the placenta until 18 weeks of gestation (Abramovich and Page, 1972), and testosterone production in males is at its highest before this gestational age (Abramovich, 1974; Nagamani et al., 1979). These facts, combined with the considerable evidence of intrauterine position in animals, suggest that there might be exposure to elevated levels of testosterone in twins with male co-twins. However, because females produce little amounts of testosterone themselves, the effect of gestating with a male co-twin is expected to be more pronounced in females than in males (Tapp et al., 2011).

3 Postnatal socialization

An alternative explanation to prenatal hormone effects is differences in postnatal socialization. Socialization effects could result from different experiences including different exposure to sex-typed toys and activities due to having a co-twin of the opposite sex vs. having one of the same sex (Cohen-Bendahan et al., 2005b; Henderson and Berenbaum, 1997; Pulkkinen et al., 2003). Thus, it is possible that OS females are raised in a more male-typical environment than SS females and that this may affect their behavior. In addition, it is also possible that the behavior of OS male twins is influenced by growing up with a twin sister (Cohen-Bendahan et al., 2005b). Studies on twins' social interactions and the parental treatment of twins are limited (Lundborg, 2008; Pulkkinen et al., 2003), but twin relationships, albeit identical twin relationships have been suggested to be some of the most unique and intimate kinds of interpersonal bonds (Neyer, 2002; Segal, 1999). For instance, a large Finnish study of 1,874 11-12 year-old twins and their 23,200 non-twin classmates reported differences in peer-assessed socio-emotional behavior between twins and singletons (Pulkkinen et al., 2003). Twins of both sexes were rated higher than singletons in adaptive behaviors, especially in socially active behavior. However, this seemed to occur mainly in the OS twins, who were rated higher than their singleton classmates in social interaction, popularity and leadership.

4 Findings from studies comparing OS and SS twins

In this special issue the overall aim is to provide an overview of published studies investigating potential differences between OS and SS twins, focusing on the TTT hypothesis. Three papers have previously reviewed the evidence on OS twins. Miller's review concluded that comparison of OS and same-sex dizygotic (SSDZ) female twins provides a reasonable model for studies of prenatal testosterone exposure (Miller, 1994). Cohen-Bendahan's review in 2005 focused on the effects of prenatal sex hormones on sex-typed behavior, and stated that there is good evidence that human sex-typed behavior is influenced by sex hormones during prenatal development with increasing evidence from the normal population (Cohen-Bendahan et al., 2005b). In 2011, Tapp et al. provided an overview of human studies of phenotypic differences in several domains between OS and SS twins and concluded that, although the evidence is inconsistent there is enough support for the TTT hypothesis to motivate further research (Tapp et al., 2011). Since the last review in 2011, several studies on different phenotypes have been conducted using the OS vs. SS twin design. This paper summarizes studies published up to April 2019 investigating potential differences between OS and SS twins on physiological, morphological, and reproductive traits (Table 1), cognitive and perceptual traits (Table 2), behavioral traits (Table 3) and other health outcomes (Table 4).

4.1 Physiological, morphological and reproductive traits

Several studies have compared OS and SS female twins regarding physiological, morphological, and reproductive traits (Table 1). Differences between OS and SS females supporting the TTT hypothesis have been found regarding tooth size (Dempsey et al., 1999; Ribeiro et al., 2013), leukocytes telomere length (Benetos et al., 2014), and brain size in nine-year-old twins, although the latter was not replicated in adult twins (Peper et al., 2009). A recent study using magnetic resonance images to study craniofacial features among eight-year old twins was not able to replicate the findings of Peper et al. regarding brain size, but found that SS females differed from all other twin groups in craniofacial morphology mainly by having a longer and wider jaw and a longer chin (Mareckova et al., 2015) (Table 1). The results of studies of anthropometric measures were largely inconsistent. No differences were identified for self-reported height (Gaist et al., 2000; Loehlin and Martin, 1998), weight, body mass index (BMI), and waist circumference (Gaist et al., 2000; Korsoff et al., 2014), but a large recent study found that BMI, body weight and the rate of dyslipidemia were moderately higher in OS than in SS females, but only among individuals age 60 and above (Alexanderson et al., 2011). A large study based on twins from 19 countries found no differences for BMI and overweight/obesity, but found that OS females were 0.31 cm taller than SS females. As the authors suggested, this difference was detected due to the large sample size used in the study and is unlikely to reflect meaningful differences of public health relevance (Bogl et al., 2017). Neither a Slovenian sample (Tul et al., 2012) nor a sample of twins from 15 countries (Jelenkovic et al., 2018) showed significant differences in birth size or gestational age (Table 1).

The Geschwind-Behan-Galaburda (GBG) hypothesis (Geschwind and Behan, 1982) postulates that high levels of testosterone may inhibit development of the left hemisphere and enhance development of the right hemisphere, resulting in increased left-handedness. In contrast to the GBG hypothesis, the callosal theory proposes that low prenatal testosterone levels result in less regressive development of temporo-parietal regions of the brain, resulting in a larger isthmus of the corpus callosum and less functional asymmetry, thus increasing left-handedness (Witelson, 1991). Thus, while the GBG hypothesis predicts a higher prevalence of left-handedness, the callosal theory predicts a lower prevalence of left-handedness among OS than SS females. Most studies on handedness have found no differences between OS and SS twins (Elkadi et al., 1999; Medland et al., 2009; Ooki, 2006), but one study found a lower prevalence of left-handedness in OS than in SS females (Vuoksimaa et al., 2010a), supporting the callosal theory of the TTT hypothesis. The second-to-fourth-finger ratio (2D:4D) is a measure of the relative length of the second finger to that of the fourth finger. The development of the ratio has been suggested to be affected by testosterone, with males having a lower 2D:4D than females on average (Manning, 2002). The four studies on finger length ratios showed conflicting results with the most recent and largest studies (Hiraishi et al., 2012; Medland et al., 2008a) finding no differences. In contrast, the two other studies (van Anders et al., 2006; Voracek and Dressler, 2007) found 2D:4D lower, e.g. more masculinized in OS than in SS females although this result was only found for the left-hand digit ratio in the Canadian study (van Anders et al., 2006) (Table 1). Some inconsistencies in findings of reproductive characteristics have been reported. A Finnish study including twins born 1734-1888 found that OS females were 25% less likely to reproduce compared with SS females (Lummaa et al. 2007). The authors suggested that the results might provide support for the TTT hypothesis that testosterone from male co-twins is associated with the impaired fertility of females. A recent population-based study from Norway supported this conclusion by finding that OS females had 11.7% lower probability of ever having being married at age 32 and that OS females had 5.8% fewer children on average (Butikofer et al., 2019) (Table 1). However, other studies within this field have not reported any differences between OS and SS twins, either for waiting time to pregnancy (Christensen et al., 1998), or for self-reported reproductive functions including age at first pregnancy and number of children (Korsoff et al., 2014; Loehlin and Martin, 1998; Medland et al., 2008b; Rose et al., 2002). One study reported slightly later menarche for OS compared with SS females (Kaprio et al., 1995), whereas more recent studies did not find any differences (Rose et al., 2002; Sorensen et al., 2013). No differences were found either for self-reported pubertal development (Rose et al., 2002) nor for the prevalence of polycystic ovary syndrome (Kuijper et al., 2009) (Tables 1 and 3). Two studies measuring testosterone (Cohen-Bendahan et al., 2004; Vuoksimaa et al., 2010a) and estradiol levels (Vuoksimaa et al., 2010a) in saliva from radioimmunoassay among children and adolescents did not find differences between OS and SS twins (Table 1).

Among studies investigating differences in physiological and morphological traits in males (Table 1), only one study investigating brain volume in nine-year-old male twins reported findings in the direction predicted by the TTT hypothesis (Peper et al., 2009). However, the results of larger total brain and cerebellum volumes for SS than for OS male children were not replicated for adult twins. Few other studies have found differences between OS and SS male twins, but not in the expected direction predicted by the TTT hypothesis (Bogl et al., 2017; Jelenkovic et al., 2018; Tul et al., 2012).

4.2 Perceptual and cognitive traits

Studies of otoacoustic emissions, which are sounds produced by the inner ear either in response to a sound (click-evoked otoacoustic emissions – CEOAEs) or in the absence of any stimulus (spontaneous otoacoustic emissions - SOAEs), have reported that SOAEs are more numerous and CEOAEs stronger in females compared with males (McFadden and Shubel, 2003; Talmadge et al., 1993). Two twin studies of predominantly the same samples found that OS females produced significantly fewer SOAEs than SS females (McFadden, 1993) and that there was a marginally significant reduction in CEOAEs for OS compared with SS females (McFadden et al., 1996), supporting the TTT hypothesis (Table 2). Also, in support of the TTT hypothesis, OS females showed a more masculine pattern of cerebral lateralization compared with SS females (Cohen-Bendahan et al., 2004) (Table 1). Moreover, studies of visuo-spatial abilities found significantly better mental rotation test (MRT) performance for OS than SS females in the direction towards the male mean (Heil et al., 2011; Vuoksimaa et al., 2010b) (Table 2). Two studies of young twins (approximately age two years) investigated expressive vocabulary by parental reports, and reported that SS females had a larger vocabulary than OS females (Galsworthy et al., 2000; Van Hulle et al., 2004), also consistent with the TTT hypothesis. However, a large Danish register study investigating academic performance measured as ninth grade test scores and teacher ratings confirmed the known sex differences from previous literature, but contrary to the hypothesis, OS females did not perform better than SS females in mathematics, nor did they perform worse in Danish or English (Ahrenfeldt et al., 2015a) (Table 2). The recent Norwegian register-based study found that OS females had 15.2% decreased probability of graduating high school and 3.9% lower probability of graduating from college than SS females. The study also showed that OS females had 8.6% lower life cycle-earning at age 32 compared with SS females (Butikofer et al., 2019) (Table 1). Interestingly, the same differences were observed among subsets of females whose male co-twins died during the first postnatal year, suggesting that the differences are due primarily to prenatal exposure rather than to postnatal socialization effects of being raised with a male co-twin.

Among studies investigating cognitive and perceptual traits for males (Table 2), the two twin studies investigating expressive vocabulary in young twins reported that OS males had a larger vocabulary than SS males, in accordance with the TTT hypothesis (Galsworthy et al., 2000; Van Hulle et al., 2004). The finding of no differences in academic performance (Ahrenfeldt et al., 2015a; Butikofer et al., 2019) agreed with the other studies investigating perceptual and cognitive traits among males, showing no differences between OS and SS male twins (McFadden, 1993; McFadden et al., 1996; Vuoksimaa et al., 2010b).

4.3 Behavioral traits

Evidence from studies of behavioral traits in OS and SS female twins comes from two large studies in which the results of sensation-seeking, including experience-seeking in OS compared with SS females, tended towards the male mean (Resnick et al., 1993; Slutske et al., 2011). Notably, another study found lower experience seeking among OS females compared with SS females among 13-year old female twins – opposite to what was predicted (Cohen-Bendahan et al., 2005a). Single studies have reported more masculinized scores for OS than SS females on measures of masculine attitudes (Miller and Martin, 1995), masculinity-femininity (Verweij et al., 2016), rule-breaking behavior (for the youngest of the two subsamples) (Loehlin and Martin, 2000) and aggression (Cohen-Bendahan et al., 2005a). In contrast, other studies did not find support for the TTT hypothesis in females in self-reported feminine interests (Rose et al., 2002), for Worried and Reserved masculinity-femininity subscales (Loehlin and Martin, 2000), for some aspects of social behavior and friendships (Laffey-Ardley and Thorpe, 2006), for hyperactivity disorder (ADHD) and autistic symptomatology (Attermann et al., 2012; Eriksson et al., 2016; Ho et al., 2005) or for religiousness (Ahrenfeldt et al., 2016). Moreover, other reports have been negative including a comparison of parental reports of play activities and playmate preferences (Elizabeth and Green, 1984) and two observational studies of sex-typed childhood play (Henderson and Berenbaum, 1997; Rodgers et al., 1998). In the latter studies, however, the sample sizes were small (Table 3). Using the same methodology, four studies investigated disordered eating. One study found that OS females reported less disordered eating than SS females (Culbert et al., 2008), but the other three studies reported no differences (Baker et al., 2009; Lydecker et al., 2012; Raevuori et al., 2008). One recent study found lower levels of disordered eating for OS than for SS females during mid-late puberty but no differences during pre-early puberty (Culbert et al., 2013). No difference in the prevalence of alcohol-dependence was found between OS and SS females in large samples from Sweden and Australia (Lenz et al., 2012), but another study from Australia found that OS females had slightly more lifetime alcohol use disorder symptoms than SS females (Ellingson et al., 2013) (Table 3).

Overall, the evidence for prenatal hormone transfer on behavioral traits for males is sparse with most studies showing no differences between OS and SS males on various sensation-seeking scores (Resnick et al., 1993), feminine interests (Rose et al., 2002) and masculine attitudes (Miller and Martin, 1995). However, one study (Ho et al., 2005) reported that sub-threshold autistic symptomatology rated by parents was higher in SS than in OS male twins. Moreover, a single study reported more disordered eating for OS than for SS males (Culbert et al., 2008). In an older adult cohort, Loehlin and Martin found that OS males scored more feminine than SS males on a worried subscale. This scale contrasts individuals who describe themselves as fearful and worried with those describing themselves as calm and confident. Women score on average higher than men. Thus, the results from the older cohort suggest that having a female co-twin makes the male co-twin more anxious (Loehlin and Martin, 2000). However, for the younger twin cohort, results for this worried scale pointed in the opposite direction, and results for rule-breaking were also in contrast to what was expected, with OS males having more masculine scores than SS males (Loehlin and Martin, 2000). For alcohol dependence, OS males were more likely to become alcohol-dependent than SS males, and this was interpreted as indirect evidence for the role of prenatal testosterone on alcohol dependence (Lenz et al., 2012). One study on sex-typed behavior based on parental reports suggested that OS males might be demasculinized on sex-typical behavior (Elizabeth and Green, 1984). However, no differences were found in an observational study on two-year old children regarding toy preferences (Rodgers et al., 1998) (Table 3).

4.4 Other health traits

Recent register-based twin studies on epilepsy (Mao et al., 2018) and hormone-related cancers (Ahrenfeldt et al., 2015b) did not show differences between OS and SS twins, either for females or for males. In addition, a recent study on early life mortality risks found that OS girls did not have higher mortality risks than SS girls. However, in line with what was hypothesized, OS boys showed lower mortality than SS boys persisting within the first year of life. This finding might provide evidence for the TTT hypothesis, but according to the authors this was at least partly due to the inclusion of MZ twins in the SS twin group (Ahrenfeldt et al., 2017) (Table 4 and chapter 6.2).

5 Summary of main findings

5.1 Recent findings in relation to the TTT hypothesis

In this special issue, we summarize published studies investigating potential differences between OS and SS female and male twins, respectively focusing on the TTT hypothesis. A total of 60 articles fulfilled the eligibility criteria including 23 studies published since the review by Tapp et al. in 2011. Among recent studies of physiological, morphological and reproductive traits in OS and SS females, three studies were interpreted as providing evidence for the TTT hypothesis (Benetos et al., 2014; Butikofer et al., 2019; Ribeiro et al., 2013), whereas six studies (Bogl et al., 2017; Hiraishi et al., 2012; Jelenkovic et al., 2018; Korsoff et al., 2014; Sorensen et al., 2013; Tul et al., 2012) did not find support for prenatal hormone transfer *in utero*, and two studies (Alexanderson et al., 2011; Mareckova et al., 2015) reported some differences between OS and SS female twins. Mareckova et al. (2015) found that SS females differed from OS females in craniofacial morphology, but did not replicate previous findings on brain size. Alexanderson et al. (2011) found differences on anthropometric measures, but the findings were limited to individuals aged 60 and above.

In general, outcomes from OS twin studies on perception (mainly otoacoustic emissions) and cognition (expressive vocabulary and visuo-spatial ability) are mainly supportive of the TTT hypothesis especially for females, but studies of these traits have not been replicated in recent years. One recent study on cognitive traits did not provide evidence for the TTT hypothesis on academic performance in adolescence (Ahrenfeldt et al., 2015a), but a recent Norwegian register-based study found that OS females had lower educational and socioeconomic attainments than SS females (Butikofer et al., 2019). Among recent studies of behavioral traits, five studies did not report differences between OS and SS females (Ahrenfeldt et al., 2016; Attermann et al., 2012; Eriksson et al., 2016; Lenz et al., 2012; Lydecker et al., 2012), whereas two studies showed some differences in the expected direction on lifetime alcohol use disorder symptoms (Ellingson et al., 2013) and disordered eating (Culbert et al., 2013). One twin study found differences regarding masculinity-femininity in the hypothesized direction for females, but in the opposite direction for males (Verweij et al., 2016). Few studies have investigated whether health outcomes (including epilepsy, hormone related cancers and early life mortality risks) differ between OS and SS twins, but recent register-based studies on larges samples of Nordic twins did not find differences between OS and SS twins in the expected directions (Ahrenfeldt et al., 2017; Ahrenfeldt et al., 2015b; Mao et al., 2018).

Among the 13 recent studies including males, only few studies have provided some evidence for the TTT hypothesis (Ahrenfeldt et al., 2017; Lenz et al., 2012). Nevertheless, although some differences between OS and SS twins have been reported for several traits, there seems to be a growing body of research challenging the manifestation of the TTT hypothesis in observational studies.

5.2 Findings in relation to postnatal socialization

A concern in interpreting findings from OS twin studies is that the effects of the social environment can be interpreted as support for the TTT hypothesis (Tapp et al., 2011). However, although it may be difficult to separate postnatal psychosocial from prenatal hormonal influences, in general, there is sparse evidence regarding twins' social interactions. Several studies of phenotypes for which psycho-social influences could be expected, for instance in toy preferences and aggression, did not show differences between OS and SS twins, whereas other domains in which social influences are expected to be minimal, such as cerebral lateralization, otoacoustic emissions and tooth size, are consistent with the TTT hypothesis. Few studies have used siblings as a control group for the psychosocial environment, but these studies have failed to find evidence of socialization effects (Ellingson et al., 2013; Heil et al., 2011; Henderson and Berenbaum, 1997; Slutske et al., 2011). In general, investigations of OS female twins have considered a hormonal, rather than a social explanation of possible sex-typing effects (Miller, 1994).

6 Methodological considerations

6.1 The value of twins for testing testosterone effects

So far, there is no direct evidence that OS females are exposed to higher testosterone levels during prenatal development than SS females. While some of the results could be interpreted as providing supporting evidence for the TTT hypothesis, it should be emphasized that hormonal exposure was not tested directly. Transfer of testosterone is assumed, based on animal studies with placentation patterns which are different from those of human twin pregnancies (Ryan and Vandenbergh, 2002). Moreover, the non-human animal studies on intrauterine position that forms the basis of OS studies find that masculinization is most likely for female fetuses that gestate between two males, with smaller effects for those positioned next to just one male (Ryan and Vandenbergh, 2002). Thus, even if hormone transfer exists between twins, exposure from one male co-twin may not be high enough to induce masculinization of the female or the effects might be counteracted by other hormones (Cohen-Bendahan et al., 2005b). In general, to make a valid investigation of the TTT hypothesis, the investigated phenotypes in OS/SS twin studies should be correlated with testosterone. However, the association with androgen exposure for some of the investigated outcomes are uncertain, for instance regarding autistic traits (Kung et al., 2016a; Kung et al., 2016b) and 2D:4D ratio (Berenbaum et al., 2009). Thus, the null findings for OS twin studies on these traits could also suggest that there is no influence of potential, excessive testosterone exposure *in utero* on these traits.

Available evidence suggests that prenatal testosterone influences human behaviors that show large differences between males and females (Hines et al., 2016), and the size of the sex difference should be considered in interpreting studies of OS and SS twins. The most common measure of the magnitude of gender differences in psychological research is Cohen's *d,* which is calculated as the mean of males minus the mean of females, divided by the pooled within-groups standard deviation (Cohen, 1988). According to Cohen's guidelines, $d = 0.20$ is interpreted as a small difference, $d = 0.50$ is a moderate difference and $d = 0.80$ is interpreted as a large difference. The most consistent sex difference in cognitive abilities favoring males exists for visuospatial abilities and is found in MRT performance with large effect sizes (Voyer et al., 1995). Among OS twin studies on cognitive function, Vuoksimaa et al. found, in a study of 804 Finnish twins that OS females outperformed SS females in MRT performance ($d = 0.30$), also significant after controlling for possible confounding variables including age, birth weight, gestational age, maternal age, and computer game experience (Vuoksimaa et al., 2010b). Heil et al. replicated the results of better MRT performance for OS than SS females ($d = 0.38$) in a sample of 200 female twins and 200 non-twin controls. They also found that OS females had higher MRT performance than females raised with a slightly older brother, which according to the authors highlighted the organizational effects of prenatal testosterone (Heil et al., 2011). MRT ability is thought to have implications for standardized exams in mathematics (Halpern et al., 2007). However, a recent register-based study on academic performance did not find differences between OS and SS twins in mathematics in the predicted direction, based on the TTT hypothesis (Ahrenfeldt et al., 2015a). This may be because the tasks on which the ninth-grade test scores and teacher ratings were based rely upon many other abilities than the MRT performance, or because sex differences in average mathematical tests are small (Else-Quest et al., 2010; Lindberg et al., 2010; Stoet and Geary, 2013). However, the size of the difference in mathematics scores ($d = 0.06$-0.15) was comparable with the sex difference in mathematics, but in the opposite direction (Ahrenfeldt et al., 2015a). In general, among OS twin studies reporting effect sizes (Cohen's d), several studies found only small differences between OS and SS twins, and some of these findings are unlikely to be of clinical relevance e.g. the findings by Bogl et al. (2017) and by Kaprio et al. (1995). As with any review or meta-analysis, the findings in this study may be influenced by publication bias. Because statistically significant results are more likely to be published than null findings (Dickersin et al., 1987), the evidence against the TTT hypothesis may be even more pronounced than suggested in this review paper.

However, a problematic issue with a number of these studies is their inability to disentangle the OS twins' prenatal hormonal environment from their postnatal social environment. In other words, finding more masculinized traits among OS female twins, relative to SS female twins, is consistent with both hormonal exposure and socialization. Two recent studies have been able to accomplish this task using unique genetically and environmentally informative kinships. One such study compared the nonverbal cognitive performance of OS female twins to that of OS virtual female twins (Segal et al., 2019). Virtual twins are same-age unrelated siblings raised together since birth, who replicate the twin relationship, but without the biological link. The sample size was small, but among the younger participants a trend toward higher scores on the Block Design subscale (a test of visuospatial ability) of the Wechsler IQ Test was found among the OS female twins who additionally shared their prenatal hormonal environment. This finding provides some support for the prenatal masculinization of these female twins. A second recent study reported reduced fertility and socioeconomic achievement among OS female twins, relative to SS female twins (Butikofer et al., 2019). More importantly, these findings were replicated using OS female twins whose cotwins had died during the first year of life, effectively separating the prenatal and postnatal environments. The addition of raised apart OS twins would bring an additional dimension to this work, because these co-twins share their prenatal circumstances, but not their rearing environment (Segal, 2012).

6.2 The monozygotic twins

Ideally, the TTT hypothesis should be tested using only DZ twins because the intrauterine environment and, thus, possible androgen exposure in twins may be affected by other factors of importance in twin pregnancies that may differ between MZ and DZ twins. They include, for example, left-right position in the uterus, placentation, number of chorions and twin differences in growth (Cohen-Bendahan et al., 2005b). Although several studies have restricted their samples to DZ twin pairs, it was not always possible to allocate all the SS twins to zygosity groups. For instance, 51% of twins in the Danish study investigating early life mortality risks (Ahrenfeldt et al., 2017) were twins with unknown zygosity (UZ). It is generally stated that twins whose zygosity is not accessible are disadvantaged compared with twins of known zygosity (Petersen et al., 2011) and that they are not representative of the twin population at large (Madsen et al., 2010). For example, UZ twins have lower socio-economic status and poorer school performance compared with the other twin groups (Petersen et al., 2009), and they have lower marriage rates and higher divorce rates compared with twins whose zygosity is known (Petersen et al., 2011). Zygosity information is available for all OS twins, i.e. twins who are less privileged are also included in this group. Thus, if the UZ twins are excluded, the SS twin group may not be comparable with the OS twins and bias may be present. It is possible, however, to exclude twins with known MZ status by comparing OS with SSDZ and UZ twins combined. Most studies have investigated DZ twins only, but most often there is no information about UZ twins in the OS twin studies. It may be preferable, however, to keep MZ twins within the comparisons if the MZ and DZ twins do not differ on the outcome variables.

7 Conclusion and future directions

The overall aim of this paper was to provide an overview of published studies investigating potential differences between OS and SS twins focusing on the TTT hypothesis. A total of 60 articles fulfilled the eligibility criteria, including 23 studies published since the last review article (Tapp et al., 2011). Overall, studies of cognition are conflicting, but cognition is the phenotype for which most support for the TTT hypothesis is found. Evidence regarding physiological and behavioral traits is less consistent. Most studies have failed to identify differences between OS and SS twin males. Since the last review, only a limited number of differences in physiological, cognitive and behavioral outcomes between OS and SS twins have been reported, and there seems to be a growing body of research challenging the manifestation of the TTT hypothesis in observational studies. Thus, although the value of twins for testing hormone effects is uncertain, findings are generally reassuring as they fail to find evidence for important differences between OS and SS twins on, for instance, health outcomes such as cancer risks later in life. We hope that this contribution to the special issue will stimulate a discussion about how an investigation of the TTT hypothesis should continue in future research.

To date, the outcome measures in the comparisons of OS and SS twins are mainly self-reported and only few studies have used registry-based outcomes. The large, worldwide twin registries linked to several health and administrative registers may be a useful tool for studying the effects of prenatal hormones as frequency of DZ twinning is increasing globally (Hoekstra et al., 2008). Nevertheless, future twin studies should take into account the role of the social environment. To do that, it would be best to examine virtual twins (unrelated siblings of the same sex raised together from early infancy) or twins reared apart (Segal, 2000; Segal et al., 2019). However, these subjects are rare. Another method would be to include singleton sibling pairs or siblings of twins (Cohen-Bendahan et al., 2005b; Henderson and Berenbaum, 1997) for instance by comparing OS females with SS females having a slightly older or younger brother. Differences between siblings may be related to the postnatal environment, contrary to twin pairs who share both a prenatal and a postnatal environment. In this way, the comparison group of siblings can help to differentiate between behavioral effects due to prenatal hormones and to the social environment, although twin pairs may affect each other differently than siblings of different ages (Cohen-Bendahan et al., 2005b). Only few studies (Ellingson et al., 2013; Heil et al., 2011; Henderson and Berenbaum, 1997; Slutske et al., 2011) have included such a comparison group. Moreover, age effects are important to consider in future studies, because prenatal testosterone effects may diminish with age as the influences of other hormonal or environmental factors become predominant (Cohen-Bendahan et al., 2005b). To study age effects, ideally longitudinal studies of OS vs. SS twins with sibling controls are needed. However, cross-sectional investigations may be helpful too. Furthermore, most twin studies examining the TTT hypothesis have been based on Caucasian samples. Given racial/ethnic differences in levels of hormone (Agurs-Collins et al., 2012; Richard et al., 2014), future twin studies should examine non-Caucasian samples, as well, to generalize the TTT effect across human populations. Lastly, a possible extension of the traditional twin study design to investigate the TTT hypothesis is to include triplets and perhaps also quadruplets because the effects of prenatal hormone transfer in animals are most pronounced for female fetuses positioned between two males. Recently, higher order (3+) multiple births and pregnancies have become increasingly common with the use of reproductive technology, and a few higher order multiple birth registries have been developed e.g. Yokoyama (2019). Thus, it would be possible in future studies to compare females positioned *in utero* between two males with those positioned between one male and one female or between two females.

Acknowledgements

This study was supported by research grants from the National Institute on Aging (NIA-PO1-AG08761, NIAP01-AG031719) and from the European Union’s Seventh Framework Programme (FP7/2007–2011) under grant agreement no 259679.

References

Abramovich, D.R., 1974. Human sexual differentiation--in utero influences. J. Obstet. Gynaecol. Br. Commonw. 81, 448-453.

Abramovich, D.R., Page, K.R., 1972. Pathways of water exchange in the fetoplacental unit at mid-pregnancy. J.Obstet.Gynaecol.Br.Commonw. 79, 1099-1102.

Agurs-Collins, T., Rohrmann, S., Sutcliffe, C., Bienstock, J.L., Monsegue, D., Akereyeni, F., Bradwin, G., Rifai, N., Pollak, M.N., Platz, E.A., 2012. Racial variation in umbilical cord blood sex steroid hormones and the insulin-like growth factor axis in African-American and white female neonates. Cancer Causes Control 23, 445-454.

Ahrenfeldt, L., Petersen, I., Johnson, W., Christensen, K., 2015a. Academic performance of opposite-sex and same-sex twins in adolescence: A Danish national cohort study. Horm. Behav. 69, 123-131.

Ahrenfeldt, L.J., Larsen, L.A., Lindahl-Jacobsen, R., Skytthe, A., Hjelmborg, J.V., Moller, S., Christensen, K., 2017. Early-life mortality risks in opposite-sex and same-sex twins: a Danish cohort study of the twin testosterone transfer hypothesis. Ann. Epidemiol. 27, 115-120.e112.

Ahrenfeldt, L.J., Lindahl-Jacobsen, R., Moller, S., Christensen, K., Hvidtjorn, D., Hvidt, N.C., 2016. Differences in Religiousness in Opposite-Sex and Same-Sex Twins in a Secular Society. Twin research and human genetics : the official journal of the International Society for Twin Studies 19, 35-46.

Ahrenfeldt, L.J., Skytthe, A., Moller, S., Czene, K., Adami, H.O., Mucci, L.A., Kaprio, J., Petersen, I., Christensen, K., Lindahl-Jacobsen, R., 2015b. Risk of sex-specific cancers in opposite-sex and same-sex twins in Denmark and Sweden. Cancer Epidemiol. Biomarkers Prev.

Alexanderson, C., Henningsson, S., Lichtenstein, P., Holmang, A., Eriksson, E., 2011. Influence of having a male twin on body mass index and risk for dyslipidemia in middle-aged and old women. Int. J. Obes. (Lond.) 35, 1466-1469.

Attermann, J., Obel, C., Bilenberg, N., Nordenbaek, C.M., Skytthe, A., Olsen, J., 2012. Traits of ADHD and autism in girls with a twin brother: a Mendelian randomization study. Eur. Child Adolesc. Psychiatry 21, 503-509.

Baker, J.H., Lichtenstein, P., Kendler, K.S., 2009. Intrauterine testosterone exposure and risk for disordered eating. Br.J.Psychiatry 194, 375-376.

Benetos, A., Dalgard, C., Labat, C., Kark, J.D., Verhulst, S., Christensen, K., Kimura, M., Horvath, K., Kyvik, K.O., Aviv, A., 2014. Sex difference in leukocyte telomere length is ablated in opposite-sex co-twins. Int. J. Epidemiol. 43, 1799-1805.

Berenbaum, S.A., Bryk, K.K., Nowak, N., Quigley, C.A., Moffat, S., 2009. Fingers as a marker of prenatal androgen exposure. Endocrinology 150, 5119-5124.

Bogl, L.H., Jelenkovic, A., Vuoksimaa, E., Ahrenfeldt, L., Pietilainen, K.H., Stazi, M.A., Fagnani, C., D'Ippolito, C., Hur, Y.M., Jeong, H.U., Silberg, J.L., Eaves, L.J., Maes, H.H., Bayasgalan, G., Narandalai, D., Cutler, T.L., Kandler, C., Jang, K.L., Christensen, K., Skytthe, A., Kyvik, K.O., Cozen, W., Hwang, A.E., Mack, T.M., Derom, C.A., Vlietinck, R.F., Nelson, T.L., Whitfield, K.E., Corley, R.P., Huibregtse, B.M., McAdams, T.A., Eley, T.C., Gregory, A.M., Krueger, R.F., McGue, M., Pahlen, S., Willemsen, G., Bartels, M., van Beijsterveldt, T., Pang, Z., Tan, Q., Zhang, D., Martin, N.G., Medland, S.E., Montgomery, G.W., Hjelmborg, J.V.B., Rebato, E., Swan, G.E., Krasnow, R., Busjahn, A., Lichtenstein, P., Oncel, S.Y., Aliev, F., Baker, L.A., Tuvblad, C., Siribaddana, S.H., Hotopf, M., Sumathipala, A., Rijsdijk, F., Magnusson, P.K.E., Pedersen, N.L., Aslan, A.K.D., Ordonana, J.R., Sanchez-Romera, J.F., Colodro-Conde, L., Duncan, G.E., Buchwald, D., Tarnoki, A.D., Tarnoki, D.L., Yokoyama, Y., Hopper, J.L., Loos, R.J.F., Boomsma, D.I., Sorensen, T.I.A., Silventoinen, K., Kaprio, J., 2017. Does the sex

of one's co-twin affect height and BMI in adulthood? A study of dizygotic adult twins from 31 cohorts. Biol. Sex Differ. 8, 14.
Butikofer, A., Figlio, D.N., Karbownik, K., Kuzawa, C.W., Salvanes, K.G., 2019. Evidence that prenatal testosterone transfer from male twins reduces the fertility and socioeconomic success of their female co-twins. Proc. Natl. Acad. Sci. U. S. A. 116, 6749-6753.
Christensen, K., Basso, O., Kyvik, K.O., Juul, S., Boldsen, J., Vaupel, J.W., Olsen, J., 1998. Fecundability of female twins. Epidemiology 9, 189-192.
Cohen-Bendahan, C.C., Buitelaar, J.K., van Goozen, S.H., Cohen-Kettenis, P.T., 2004. Prenatal exposure to testosterone and functional cerebral lateralization: a study in same-sex and opposite-sex twin girls. Psychoneuroendocrinology 29, 911-916.
Cohen-Bendahan, C.C., Buitelaar, J.K., van Goozen, S.H., Orlebeke, J.F., Cohen-Kettenis, P.T., 2005a. Is there an effect of prenatal testosterone on aggression and other behavioral traits? A study comparing same-sex and opposite-sex twin girls. Horm.Behav. 47, 230-237.
Cohen-Bendahan, C.C., van de Beek, C., Berenbaum, S.A., 2005b. Prenatal sex hormone effects on child and adult sex-typed behavior: methods and findings. Neurosci.Biobehav.Rev. 29, 353-384.
Cohen, J., 1988. Statistical power analysis for the behavioral sciences 2nd ed. Hillsdale, NJ; Erlbaum.
Collaer, M.L., Hines, M., 1995. Human behavioral sex differences: a role for gonadal hormones during early development? Psychol. Bull. 118, 55-107.
Constantinescu, M., Hines, M., 2012. Relating Prenatal Testosterone Exposure to Postnatal Behavior in Typically Developing Children: Methods and Findings. Child Development Perspectives 6, 407-413.
Culbert, K.M., Breedlove, S.M., Burt, S.A., Klump, K.L., 2008. Prenatal hormone exposure and risk for eating disorders: a comparison of opposite-sex and same-sex twins. Arch.Gen.Psychiatry 65, 329-336.
Culbert, K.M., Breedlove, S.M., Sisk, C.L., Burt, S.A., Klump, K.L., 2013. The emergence of sex differences in risk for disordered eating attitudes during puberty: a role for prenatal testosterone exposure. J. Abnorm. Psychol. 122, 420-432.
Dempsey, P.J., Townsend, G.C., Richards, L.C., 1999. Increased tooth crown size in females with twin brothers: Evidence for hormonal diffusion between human twins in utero. Am.J.Hum.Biol. 11, 577-586.
Dickersin, K., Chan, S., Chalmers, T.C., Sacks, H.S., Smith, H., Jr., 1987. Publication bias and clinical trials. Control. Clin. Trials 8, 343-353.
Elizabeth, P.H., Green, R., 1984. Childhood sex-role behaviors: similarities and differences in twins. Acta Genet.Med.Gemellol.(Roma.) 33, 173-179.
Elkadi, S., Nicholls, M.E., Clode, D., 1999. Handedness in opposite and same-sex dizygotic twins: testing the testosterone hypothesis. Neuroreport 10, 333-336.
Ellingson, J.M., Slutske, W.S., Richmond-Rakerd, L.S., Martin, N.G., 2013. Investigating the influence of prenatal androgen exposure and sibling effects on alcohol use and alcohol use disorder in females from opposite-sex twin pairs. Alcohol. Clin. Exp. Res. 37, 868-876.
Else-Quest, N.M., Hyde, J.S., Linn, M.C., 2010. Cross-national patterns of gender differences in mathematics: a meta-analysis. Psychol. Bull. 136, 103-127.
Eriksson, J.M., Lundstrom, S., Lichtenstein, P., Bejerot, S., Eriksson, E., 2016. Effect of co-twin gender on neurodevelopmental symptoms: a twin register study. Mol. Autism 7, 8.
Fishman, R., Vortman, Y., Shanas, U., Koren, L., 2019. Non-model species deliver a non-model result: Nutria female fetuses neighboring males in utero have lower testosterone. Horm. Behav.
Gaist, D., Bathum, L., Skytthe, A., Jensen, T.K., McGue, M., Vaupel, J.W., Christensen, K., 2000. Strength and anthropometric measures in identical and fraternal twins: no evidence of masculinization of females with male co-twins. Epidemiology 11, 340-343.
Galsworthy, M.J., Plomin, R., Dionne, G., Dale, P.S., 2000. Sex differences in early verbal and non-verbal cognitive development. Developmental Science 3, 206.
Geschwind, N., Behan, P., 1982. Left-handedness: association with immune disease, migraine, and developmental learning disorder. Proc. Natl. Acad. Sci. U. S. A. 79, 5097-5100.

Glass, A.R., Klein, T., 1981. Changes in maternal serum total and free androgen levels in early pregnancy: lack of correlation with fetal sex. Am. J. Obstet. Gynecol. 140, 656-660.
Hall, J.G., 2003. Twinning. Lancet 362, 735-743.
Halpern, D.F., Benbow, C.P., Geary, D.C., Gur, R.C., Hyde, J.S., Gernsbacher, M.A., 2007. The Science of Sex Differences in Science and Mathematics. Psychol. Sci. Public Interest 8, 1-51.
Hampson, E., Rovet, J.F., Altmann, D., 1998. Spatial reasoning in children with congenital adrenal hyperplasia due to 21-hydroxylase deficiency. Dev. Neuropsychol. 14, 299-320.
Heil, M., Kavsek, M., Rolke, B., Beste, C., Jansen, P., 2011. Mental rotation in female fraternal twins: Evidence for intra-uterine hormone transfer? Biol.Psychol. 86, 90-93.
Henderson, B.A., Berenbaum, S.A., 1997. Sex-typed play in opposite-sex twins. Dev.Psychobiol. 31, 115-123.
Hickey, M., Sloboda, D.M., Atkinson, H.C., Doherty, D.A., Franks, S., Norman, R.J., Newnham, J.P., Hart, R., 2009. The relationship between maternal and umbilical cord androgen levels and polycystic ovary syndrome in adolescence: a prospective cohort study. J. Clin. Endocrinol. Metab. 94, 3714-3720.
Hines, M., 2006. Prenatal testosterone and gender-related behaviour. Eur. J. Endocrinol. 155 Suppl 1, S115-121.
Hines, M., 2008. Early androgen influences on human neural and behavioural development. Early Hum. Dev. 84, 805-807.
Hines, M., 2010. Sex-related variation in human behavior and the brain. Trends Cogn. Sci. 14, 448-456.
Hines, M., 2011. Gender development and the human brain. Annu. Rev. Neurosci. 34, 69-88.
Hines, M., Fane, B.A., Pasterski, V.L., Mathews, G.A., Conway, G.S., Brook, C., 2003. Spatial abilities following prenatal androgen abnormality: targeting and mental rotations performance in individuals with congenital adrenal hyperplasia. Psychoneuroendocrinology 28, 1010-1026.
Hines, M., Golombok, S., Rust, J., Johnston, K.J., Golding, J., 2002. Testosterone during pregnancy and gender role behavior of preschool children: a longitudinal, population study. Child Dev. 73, 1678-1687.
Hines, M., Spencer, D., Kung, K.T., Browne, W.V., Constantinescu, M., Noorderhaven, R.M., 2016. The early postnatal period, mini-puberty, provides a window on the role of testosterone in human neurobehavioural development. Curr. Opin. Neurobiol. 38, 69-73.
Hiraishi, K., Sasaki, S., Shikishima, C., Ando, J., 2012. The second to fourth digit ratio (2D:4D) in a Japanese twin sample: heritability, prenatal hormone transfer, and association with sexual orientation. Arch. Sex. Behav. 41, 711-724.
Ho, A., Todd, R.D., Constantino, J.N., 2005. Brief report: autistic traits in twins vs. non-twins--a preliminary study. J. Autism Dev. Disord. 35, 129-133.
Hoekstra, C., Zhao, Z.Z., Lambalk, C.B., Willemsen, G., Martin, N.G., Boomsma, D.I., Montgomery, G.W., 2008. Dizygotic twinning. Hum.Reprod.Update. 14, 37-47.
Jelenkovic, A., Sund, R., Yokoyama, Y., Hur, Y.M., Ullemar, V., Almqvist, C., Magnusson, P.K., Willemsen, G., Bartels, M., Beijsterveldt, C.E.V., Bogl, L.H., Pietilainen, K.H., Vuoksimaa, E., Ji, F., Ning, F., Pang, Z., Nelson, T.L., Whitfield, K.E., Rebato, E., Llewellyn, C.H., Fisher, A., Bayasgalan, G., Narandalai, D., Bjerregaard-Andersen, M., Beck-Nielsen, H., Sodemann, M., Tarnoki, A.D., Tarnoki, D.L., Ooki, S., Stazi, M.A., Fagnani, C., Brescianini, S., Dubois, L., Boivin, M., Brendgen, M., Dionne, G., Vitaro, F., Cutler, T.L., Hopper, J.L., Krueger, R.F., McGue, M., Pahlen, S., Craig, J.M., Saffery, R., Haworth, C.M., Plomin, R., Knafo-Noam, A., Mankuta, D., Abramson, L., Burt, S.A., Klump, K.L., Vlietinck, R.F., Derom, C.A., Loos, R.J., Boomsma, D.I., Sorensen, T.I.A., Kaprio, J., Silventoinen, K., 2018. Birth size and gestational age in opposite-sex twins as compared to same-sex twins: An individual-based pooled analysis of 21 cohorts. Sci. Rep. 8, 6300.
Kaprio, J., Rimpela, A., Winter, T., Viken, R.J., Rimpela, M., Rose, R.J., 1995. Common genetic influences on BMI and age at menarche. Hum. Biol. 67, 739-753.
Korsoff, P., Bogl, L.H., Korhonen, P., Kangas, A.J., Soininen, P., Ala-Korpela, M., Rose, R.J., Kaaja, R., Kaprio, J., 2014. A comparison of anthropometric, metabolic, and reproductive characteristics of young adult women from opposite-sex and same-sex twin pairs. Front. Endocrinol. (Lausanne) 5, 28.
Kuijper, E.A., Vink, J.M., Lambalk, C.B., Boomsma, D.I., 2009. Prevalence of polycystic ovary syndrome in women from opposite-sex twin pairs. J.Clin.Endocrinol.Metab 94, 1987-1990.

Kung, K.T., Constantinescu, M., Browne, W.V., Noorderhaven, R.M., Hines, M., 2016a. No relationship between early postnatal testosterone concentrations and autistic traits in 18 to 30-month-old children. Mol. Autism 7, 15.
Kung, K.T., Spencer, D., Pasterski, V., Neufeld, S., Glover, V., O'Connor, T.G., Hindmarsh, P.C., Hughes, I.A., Acerini, C.L., Hines, M., 2016b. No relationship between prenatal androgen exposure and autistic traits: convergent evidence from studies of children with congenital adrenal hyperplasia and of amniotic testosterone concentrations in typically developing children. J. Child Psychol. Psychiatry 57, 1455-1462.
Laffey-Ardley, S., Thorpe, K., 2006. Being opposite: is there advantage for social competence and friendships in being an opposite-sex twin? Twin research and human genetics : the official journal of the International Society for Twin Studies 9, 131-140.
Lenz, B., Muller, C.P., Kornhuber, J., 2012. Alcohol dependence in same-sex and opposite-sex twins. J. Neural Transm. 119, 1561-1564.
Lewi, L., 2010. Monochorionic diamniotic twin pregnancies pregnancy outcome, risk stratification and lessons learnt from placental examination. Verh. K. Acad. Geneeskd. Belg. 72, 5-15.
Lindberg, S.M., Hyde, J.S., Petersen, J.L., Linn, M.C., 2010. New trends in gender and mathematics performance: a meta-analysis. Psychol. Bull. 136, 1123-1135.
Loehlin, J.C., Martin, N.G., 1998. A comparison of adult female twins from opposite-sex and same-sex pairs on variables related to reproduction. Behav.Genet. 28, 21-27.
Loehlin, J.C., Martin, N.G., 2000. Dimensions of psychological masculinity-femininity in adult twins from opposite-sex and same-sex pairs. Behav.Genet. 30, 19-28.
Loos, R., Derom, C., Vlietinck, R., Derom, R., 1998. The East Flanders Prospective Twin Survey (Belgium): a population-based register. Twin Res. 1, 167-175.
Lundborg, P., 2008. The Health Returns to Education: What can we learn from twins?, pp. 1-32.
Lydecker, J.A., Pisetsky, E.M., Mitchell, K.S., Thornton, L.M., Kendler, K.S., Reichborn-Kjennerud, T., Lichtenstein, P., Bulik, C.M., Mazzeo, S.E., 2012. Association between co-twin sex and eating disorders in opposite sex twin pairs: evaluations in North American, Norwegian, and Swedish samples. J. Psychosom. Res. 72, 73-77.
Madsen, M., Andersen, A.M., Christensen, K., Andersen, P.K., Osler, M., 2010. Does educational status impact adult mortality in Denmark? A twin approach. Am.J.Epidemiol. 172, 225-234.
Manning, J.T., 2002. Digit ratio: a pointer to fertility, behavior and health. Rutgers University Press, New Brunswick, New Jersey, and London.
Mao, Y., Ahrenfeldt, L.J., Christensen, K., Wu, C., Christensen, J., Olsen, J., Sun, Y., 2018. Risk of epilepsy in opposite-sex and same-sex twins: a twin cohort study. Biol. Sex Differ. 9, 21.
Mareckova, K., Chakravarty, M.M., Lawrence, C., Leonard, G., Perusse, D., Perron, M., Pike, B.G., Richer, L., Veillette, S., Pausova, Z., Paus, T., 2015. Identifying craniofacial features associated with prenatal exposure to androgens and testing their relationship with brain development. Brain Struct. Funct. 220, 3233-3244.
McFadden, D., 1993. A masculinizing effect on the auditory systems of human females having male co-twins. Proc.Natl.Acad.Sci.U.S.A 90, 11900-11904.
McFadden, D., Loehlin, J.C., Pasanen, E.G., 1996. Additional findings on heritability and prenatal masculinization of cochlear mechanisms: click-evoked otoacoustic emissions. Hear. Res. 97, 102-119.
McFadden, D., Shubel, E., 2003. The relationships between otoacoustic emissions and relative lengths of fingers and toes in humans. Horm. Behav. 43, 421-429.
Medland, S.E., Duffy, D.L., Wright, M.J., Geffen, G.M., Hay, D.A., Levy, F., van-Beijsterveldt, C.E., Willemsen, G., Townsend, G.C., White, V., Hewitt, A.W., Mackey, D.A., Bailey, J.M., Slutske, W.S., Nyholt, D.R., Treloar, S.A., Martin, N.G., Boomsma, D.I., 2009. Genetic influences on handedness: data from 25,732 Australian and Dutch twin families. Neuropsychologia 47, 330-337.
Medland, S.E., Loehlin, J.C., Martin, N.G., 2008a. No effects of prenatal hormone transfer on digit ratio in a large sample of same- and opposite-sex dizygotic twins. Pers. Individ. Dif. 44, 1225-1234.

Medland, S.E., Loehlin, J.C., Willemsen, G., Hatemi, P.K., Keller, M.C., Boomsma, D.I., Eaves, L.J., Martin, N.G., 2008b. Males do not reduce the fitness of their female co-twins in contemporary samples. Twin.Res.Hum.Genet. 11, 481-487.
Miller, E.M., 1994. Prenatal sex hormone transfer: A reason to study opposite-sex twins. Pers. Individ. Dif. 17, 511-529.
Miller, E.M., Martin, N., 1995. Analysis of the effect of hormones on opposite-sex twin attitudes. Acta Genet.Med.Gemellol.(Roma.) 44, 41-52.
Nagamani, M., McDonough, P.G., Ellegood, J.O., Mahesh, V.B., 1979. Maternal and amniotic fluid steroids throughout human pregnancy. Am. J. Obstet. Gynecol. 134, 674-680.
Neyer, F.J., 2002. Twin Relationships in Old Age: A Developmental Perspective. Journal of Social and Personal Relationships 19, 155-177.
Ooki, S., 2006. Nongenetic factors associated with human handedness and footedness in Japanese twin children. Environ.Health Prev.Med. 11, 304-312.
Pang, S., Levine, L.S., Cederqvist, L.L., Fuentes, M., Riccardi, V.M., Holcombe, J.H., Nitowsky, H.M., Sachs, G., Anderson, C.E., Duchon, M.A., Owens, R., Merkatz, I., New, M.I., 1980. Amniotic fluid concentrations of delta 5 and delta 4 steroids in fetuses with congenital adrenal hyperplasia due to 21 hydroxylase deficiency and in anencephalic fetuses. J. Clin. Endocrinol. Metab. 51, 223-229.
Peper, J.S., Brouwer, R.M., Van Baal, G.C., Schnack, H.G., van, L.M., Boomsma, D.I., Kahn, R.S., Hulshoff Pol, H.E., 2009. Does having a twin brother make for a bigger brain? Eur.J.Endocrinol. 160, 739-746.
Petersen, I., Jensen, V.M., McGue, M., Bingley, P., Christensen, K., 2009. No evidence of genetic mediation in the association between birthweight and academic performance in 2,413 Danish adolescent twin pairs. Twin.Res.Hum.Genet. 12, 564-572.
Petersen, I., Martinussen, T., McGue, M., Bingley, P., Christensen, K., 2011. Lower marriage and divorce rates among twins than among singletons in Danish birth cohorts 1940-1964. Twin.Res.Hum.Genet. 14, 150-157.
Phoenix, C.H., 2009. Organizing action of prenatally administered testosterone propionate on the tissues mediating mating behavior in the female guinea pig. Hormones and behavior 55, 566.
Phoenix, C.H., Goy, R.W., Gerall, A.A., Young, W.C., 1959. Organizing action of prenatally administered testosterone propionate on the tissues mediating mating behavior in the female guinea pig. Endocrinology 65, 369-382.
Pulkkinen, L., Vaalamo, I., Hietala, R., Kaprio, J., Rose, R.J., 2003. Peer reports of adaptive behavior in twins and singletons: is twinship a risk or an advantage? Twin.Res. 6, 106-118.
Raevuori, A., Kaprio, J., Hoek, H.W., Sihvola, E., Rissanen, A., Keski-Rahkonen, A., 2008. Anorexia and bulimia nervosa in same-sex and opposite-sex twins: lack of association with twin type in a nationwide study of Finnish twins. Am.J.Psychiatry 165, 1604-1610.
Resnick, S.M., Gottesman, I.I., McGue, M., 1993. Sensation seeking in opposite-sex twins: an effect of prenatal hormones? Behav.Genet. 23, 323-329.
Ribeiro, D.C., Brook, A.H., Hughes, T.E., Sampson, W.J., Townsend, G.C., 2013. Intrauterine Hormone Effects on Tooth Dimensions. J. Dent. Res.
Richard, A., Rohrmann, S., Zhang, L., Eichholzer, M., Basaria, S., Selvin, E., Dobs, A.S., Kanarek, N., Menke, A., Nelson, W.G., Platz, E.A., 2014. Racial variation in sex steroid hormone concentration in black and white men: a meta-analysis. Andrology 2, 428-435.
Rodgers, C., Fagot, B., Winebarger, A., 1998. Gender-Typed Toy Play in Dizygotic Twin Pairs: A Test of Hormone Transfer Theory. Sex Roles 39, 173-184.
Rose, R.J., Kaprio, J., Winter, T., Dick, D.M., Viken, R.J., Pulkkinen, L., Koskenvuo, M., 2002. Femininity and fertility in sisters with twin brothers: prenatal androgenization? Cross-sex socialization? Psychol.Sci. 13, 263-267.
Ryan, B.C., Vandenbergh, J.G., 2002. Intrauterine position effects. Neurosci.Biobehav.Rev. 26, 665-678.
Segal, N., 1999. Entwined Lives: Twins and what they tell us about Human Behavior. Twin Res. 2, 291-292.
Segal, N.L., 2000. Virtual twins: New findings on within-family environmental influences on intelligence. J. Educ. Psychol. 92, 442-448.

Segal, N.L., 2012. Born together-reared apart: The landmark Minnesota twin study. Harvard University Press, Cambridge, MA, USA.
Segal, N.L., Orozco, E.N., Preston, K.J.S., Gerkens, D.R., 2019. Nonverbal cognition across genetic and environmentally informative kinships: The case for opposite-sex twin pairs and virtual twins. J. Exp. Child Psychol. 182, 144-150.
Slutske, W.S., Bascom, E.N., Meier, M.H., Medland, S.E., Martin, N.G., 2011. Sensation seeking in females from opposite- versus same-sex twin pairs: hormone transfer or sibling imitation? Behav.Genet. 41, 533-542.
Sorensen, K., Juul, A., Christensen, K., Skytthe, A., Scheike, T., Kold Jensen, T., 2013. Birth size and age at menarche: a twin perspective. Hum. Reprod. 28, 2865-2871.
Stoet, G., Geary, D.C., 2013. Sex differences in mathematics and reading achievement are inversely related: within- and across-nation assessment of 10 years of PISA data. PLoS One 8, e57988.
Talmadge, C.L., Long, G.R., Murphy, W.J., Tubis, A., 1993. New off-line method for detecting spontaneous otoacoustic emissions in human subjects. Hear. Res. 71, 170-182.
Tapp, A., Maybery, M.T., Whitehouse, A.J., 2011. Evaluating the twin testosterone transfer hypothesis: A review of the empirical evidence. Horm.Behav.
Tul, N., Lucovnik, M., Novak, Z., Verdenik, I., Blickstein, I., 2012. No "masculinization" effect of a male on birth weight of its female co-twin. J. Perinat. Med. 40, 255-257.
van Anders, S.M., Vernon, P.A., Wilbur, C.J., 2006. Finger-length ratios show evidence of prenatal hormone-transfer between opposite-sex twins. Horm.Behav. 49, 315-319.
van de Beek, C., Thijssen, J.H., Cohen-Kettenis, P.T., van Goozen, S.H., Buitelaar, J.K., 2004. Relationships between sex hormones assessed in amniotic fluid, and maternal and umbilical cord serum: what is the best source of information to investigate the effects of fetal hormonal exposure? Horm. Behav. 46, 663-669.
Van Hulle, C.A., Goldsmith, H.H., Lemery, K.S., 2004. Genetic, environmental, and gender effects on individual differences in toddler expressive language. J. Speech. Lang. Hear. Res. 47, 904-912.
Verweij, K.J., Mosing, M.A., Ullen, F., Madison, G., 2016. Individual Differences in Personality Masculinity-Femininity: Examining the Effects of Genes, Environment, and Prenatal Hormone Transfer. Twin research and human genetics : the official journal of the International Society for Twin Studies 19, 87-96.
vom Saal, F.S., 1989. Sexual differentiation in litter-bearing mammals: influence of sex of adjacent fetuses in utero. J.Anim Sci. 67, 1824-1840.
Voracek, M., Dressler, S.G., 2007. Digit ratio (2D:4D) in twins: heritability estimates and evidence for a masculinized trait expression in women from opposite-sex pairs. Psychol.Rep. 100, 115-126.
Voyer, D., Voyer, S., Bryden, M.P., 1995. Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. Psychol. Bull. 117, 250-270.
Vuoksimaa, E., Eriksson, C.J., Pulkkinen, L., Rose, R.J., Kaprio, J., 2010a. Decreased prevalence of left-handedness among females with male co-twins: evidence suggesting prenatal testosterone transfer in humans? Psychoneuroendocrinology 35, 1462-1472.
Vuoksimaa, E., Kaprio, J., Kremen, W.S., Hokkanen, L., Viken, R.J., Tuulio-Henriksson, A., Rose, R.J., 2010b. Having a male co-twin masculinizes mental rotation performance in females. Psychol.Sci. 21, 1069-1071.
Warne, G.L., Faiman, C., Reyes, F.I., Winter, J.S., 1977. Studies on human sexual development. V. Concentrations of testosterone, 17-hydroxyprogesterone and progesterone in human amniotic fluid throughout gestation. J. Clin. Endocrinol. Metab. 44, 934-938.
Whitehouse, A.J.O., Maybery, M.T., Hart, R., Mattes, E., Newnham, J.P., Sloboda, D.M., Stanley, F.J., Hickey, M., 2010. Fetal androgen exposure and pragmatic language ability of girls in middle childhood: Implications for the extreme male-brain theory of autism. Psychoneuroendocrinology 35, 1259-1264.
Wilson, J.D., George, F.W., Griffin, J.E., 1981. The hormonal control of sexual development. Science 211, 1278-1284.
Witelson, S.F., 1991. Neural sexual mosaicism: sexual differentiation of the human temporo-parietal region for functional asymmetry. Psychoneuroendocrinology 16, 131-153.
Yokoyama, Y., 2019. The New West Japan Twins and Higher Order Multiple Births Registry. Twin Research and Human Genetics (in press).

Tables

Table 1 - Studies on physiological, morphological and reproductive traits in opposite-sex and same-sex twins

Trait	Publication	Setting	Numbers of opposite-sex (OS) and same-sex (SS) twins	Is there a comparison between OS and SS dizygotic (DZ) twins only?[a]	Age	Assessment	Findings OS vs. SS females	Findings OS vs. SS males	Suggests evidence for masculinization of OS females[b]	Suggests evidence for demasculinization of OS males[b]
Age at menarche	Kaprio et al., 1995	Finland	434 OSF 378 SSF (DZ) 468 SSF (MZ)	Yes	16 years	Self-reported age at menarche	OSF higher mean age at menarche than SSF (DZ) but not when compared with SSF (MZ)	NA	+	NA
	Sorensen et al., 2013	Denmark	1147 OSF 1250 SSF (DZ) 1466 SSF (MZ)	Yes	12-22 years	Self-reported age at menarche	No significant differences	NA	-	NA
Anthropometric measures	Gaist et al., 2000	Denmark	700 OSF 655 SSF (DZ) 703 SSF (MZ) 722 OSM 709 SSM (DZ) 734 SSM (MZ)	Yes	Mean = 56.9 years	Self-reported weight and height, waist circumference and handgrip strength measured by interviewers	No significant differences	No significant differences	-	-
Anthropometric measures and metabolic aberrations	Alexanderson et al., 2011	Sweden	8,409 OSF 9,166 SSF	DZ twins only	OSF: 42-93 years, mean = 59 years SSF: 42-103 years, mean = 61 years	Self-reported body mass index (BMI), hyperlipidemia and diabetes mellitus type II	BMI, body weight and rate of dyslipidemia moderately higher in OSF than in SSF but only in those over 60 years of age	NA	+/-	NA
Anthropometric measures, disease status and reproductive history	Korsoff et al., 2014	Finland	Sample A: 789 OSF 679 SSF Sample B (subsample for serum lipid and lipoprotein subclass concentrations):	DZ twins only	A: 32-37 years, mean = 34.0 years B: 21-29, mean = 24 years	Obesity measures: BMI from self-reported height and weight and waist circumference. Disease status: self-reported	No significant differences	NA	-	NA

			169 OSF 226 SSF			hypertension and diabetes. Serum lipid and lipoprotein subclass concentrations measured by nuclear magnetic resonance spectroscopy. Self-reported reproductive history (e.g. current pregnancy, number of children, abortions)				
Anthropome-tric measures	Bogl et al., 2017	CODATwins 19 countries	27,100 OSF 39,856 SSF 26,708 OSM 28,638 SSM	DZ twins only	20 years and older Males: median age = 44 years Females: median age = 42 years	Height and weight mainly self-reported (97%), BMI calculated based on weight and height	No significant differences for BMI and overweight/ obesity but OSF 0.31 cm taller than SSF	No significant differences for BMI and overweight/ obesity but OSM 0.14 cm taller than SSM	-	-
Birth size, birth weight and gestational age	Tul et al., 2012	Slovenia	1097 OSF 1554 SSF 1097 OSM 1638 SSM	No	Infants	Mean birth weight obtained from the Slovenian national information perinatal system	No significant differences	OSM heavier than SSM	-	-
	Jelenkovic et al., 2018	CODATwins 15 countries	16,417 OSF 17,432 SSF 16,417 OSM 17,584 SSM	DZ twins only	Infants	Mean birth weight, birth size and gestational age – mainly parentally reported	No significant differences	OSM on average 31 g heavier and 0.16 cm longer than SSM. OSM longer gestation than SSM	-	-
Brain size	Peper et al., 2009	The Netherlands	Sample A (children): 19 OSF 41 SSF 16 OSM 43 SSM	DZ twins only	A: mean = 9.2 years B: mean = 30.0 years	Magnetic resonance imaging (MRI) brain scans	Larger total brain and cerebellum volumes for OSF than SSF children. No	Larger total brain and cerebellum volumes for SSM than OSM children. No	+/-	+/-

			Sample B (adults): 21 OSF 40 SSF 11 OSM 44 SSM				significant differences in adults	significant differences in adults		
Brain size and craniofacial morphology	Marečková et al., 2015	Quebec, Canada	35 OSF 28 SSF 36 OSM 20 SSM	DZ twins only	8 years	Magnetic resonance images (MRI)	Brain size: No significant differences Principal components: 5/8 PC3-related distances smaller for SSF than OSF	No significant differences	+/-	-
Circulating hormone levels/ Cerebral lateralization (right ear advantage)	Cohen-Bendahan et al. 2004	The Netherlands	67 OSF 53 SSF	DZ twins only	10.4-11.8 years, mean = 11.0 years	Testosterone levels in saliva from radio-immunoassay. Auditory-verbal dichotic listening task	No significant differences regarding hormone levels. OSF greater right ear advantage of cerebral lateralization (more masculinized) than SSF	NA	+/-	NA
Circulating hormone levels/ Handedness	Vuoksimaa et al., 2010a	Finland	Testosterone/ estradiol levels 117/114 OSF 254/244 SSF 109/108 OSM 291/278 SSM Handedness 755 OSF 749 SSF (MZ) 706 SSF (DZ) 730 OSM 697 SSM (MZ) 784 SSM (DZ)	No	14 years	Testosterone and estradiol levels in saliva from radio-immunoassay. Self-reported handedness (based on two questions)	No significant differences regarding hormone levels. Lower prevalence of left-handedness in OSF than in SSF	No significant differences	+/-	-
Finger-length ratios (second to fourth finger; 2D:4D)	Van Anders et al., 2006	Canada	9 OSF 16 SSF 9 OSM 22 SSM	DZ twins only	4-15 years, mean = 10.2 years	2D:4D measures from photocopies of hands	2D:4D lower (more masculinized) in OSF than in SSF for the left-hand digit ratio	No significant differences	+	-

							but not the right			
	Voracek and Dressler, 2007	Australia	10 OSF 18 SSF	DZ twins only	Adults	2D:4D measures from photocopies of hands	2D:4D lower (more masculinized) in OSF than in SSF	NA	+	NA
	Medland et al., 2008	Australia	212 OSF 237 SSF 199 OSM 219 SSM	DZ twins only	11-24 years, mean = 15.5 years	2D:4D measures from photocopies of hands	No significant differences	No significant differences	-	-
	Hiraishi et al., 2012	Japan	9 OSF 150 SSF (MZ) 44 SSF (DZ) 9 OSM 58 SSM (MZ) 28 SSM (DZ)	No	16-32 years, mean = 22.5 years	2D:4D measures from photocopies of hands	No significant differences	No significant differences	-	-
Handedness	Elkadi et al., 1999	Australia	59 OSF 40 SSF 59 OSM 21 SSM	DZ twins only	18-26 years, OS: mean = 24.4 years SSF: mean = 24.7 years SSM: mean = 21.6 years	Edinburgh Handedness Inventory questionnaire (preferred hand for 10 everyday activities)	No significant differences	No significant differences	-	-
Handedness/ Footedness	Ooki, 2006	Japan	Handedness Sample A: 125 OSF 138 SSF (DZ) 125 OSM 150 SSM (DZ) Sample B: 203 OSF 209 SSF (DZ) 203 OSM 182 SSM (DZ) Footedness Sample A: 113 OSF 130 SSF (DZ) 114 OSM 144 SSM (DZ) Sample B: 175 OSF 172 SSF (DZ) 177 OSM 150 SSM (DZ)	Yes	A: 11-12 years B: 1-15 years, mean = 5.9 years	Self-reported hand use for writing and foot use for kicking a ball (single question)	No significant differences in prevalence except for footedness for females in sample A	No significant differences	+/-	-

Handedness	Medland et al., 2009	Australia and the Netherlands	54,270 twins and their non-twin siblings (no breakdown provided)	Yes	Children and adults	Handedness assessed by writing/ drawing or self-report	No significant differences	No significant differences	-	-
Leukocyte telomere length (LTL)	Benetos et al., 2014	Denmark	72 OSF 196 SSF (MZ) 176 SSF (DZ) 72 OSM 172 SSM (MZ) 132 SSM (DZ)	Yes	Mean (baseline) = 36.0-39.8 years	LTL measured by Southern blots of the terminal restriction fragments	Shorter LTL in OSF than in SSF	No significant differences	+	-
Polycystic ovary syndrome (PCOS)	Kuijper et al., 2009	The Netherlands	480 OSF 711 SSF (DZ) 1325 SSF (MZ)	Yes	Adults	Self-reported PCOS based	No significant differences	NA	-	NA
Reproductive function	Christensen et al., 1998	Denmark	4269 OSF 3650 SSF (MZ) 4455 SSF (DZ)	No, but differences tested between zygosity groups	18 years or older	Self-reported outcome of first try ever to become pregnant and waiting time to pregnancy	No significant differences	NA	-	NA
Reproductive function/ Height	Loehlin and Martin, 1998	Australia	Reproductive function: 600-700 OSF 1400-1500 SSF Height 685 OSF 1642 SSF	No	Reproductive function: OSF: mean = 42.6 years SSF: mean = 39.7 years Height: OSF: mean = 39.7 years SSF: mean = 42.6 years	Self-reported 90-item questionnaire related to reproductive functions and self-reported height	No significant differences	NA	-	NA
	Lummaa et al., 2007	Finland	31 OSF 35 SSF 17 OSM 26 SSM	No	Adults	Number of offspring obtained from Finnish Church registers	OSF had fewer offspring compared with SSF	No significant differences	+	-
	Medland et al., 2008b	Australia, the Netherlands and Virginia USA	913 OSF 1979 SSF	No	Adults	Self-reported number of children and age at first pregnancy	No significant differences	NA	-	NA
Reproductive function, marriage, graduating from high	Bütikofer et al., 2019	Norway	13,717 twins including 6,808 female twins (breakdown not provided)	No	At least one twin observed to age 32	Objective information on birth, household composition,	OSF decreased probability of graduating from high school (15.2%),	No significant differences	+	-

school/completing college and socioeconomic outcomes						schooling and labor market records from Norwegian Register Data	completing college (3.9%) and being married (11.7%). OSF have lower fertility (5.8%) and lower life-cycle earnings (8.6%) than SSF			
Tooth size	Dempsey et al., 1999	Australia	56 OSF 98 SSF (DZ) 166 SSF (MZ) 56 OSM 88 SSF (DZ) 132 SSF (MZ)	No	7-62 years, mean = 16.5 years	Tooth crown diameters of 28 permanent teeth recorded	OSF had larger tooth crown size than SSF for 26/28 teeth	No significant differences	+	-
	Ribeiro et al., 2013	Australia	43 OSF 39 SSF (DZ) 52 SSF (MZ)	Yes	4-16 years	Tooth crown size determined by image analysis	Larger tooth crown size in OSF than in SSF	NA	+	NA

NA: Not available, OS: opposite-sex, SS = same-sex, F = females, M = males, MZ = monozygotic, DZ = dizygotic, UZ = unknown zygosity
[a]Yes: The study makes a separate comparison of OS vs SSDZ twins, although it includes both MZ and DZ twins.
No: The study does not make a separate comparison of OS vs SSDZ twins, thus, including MZ twins in all analyses.
DZ twins only: The study includes DZ twins only. Thus, no MZ twins are included in the study.
[b]+: The study provides evidence for masculinization of OS females/demasculinization of OS males.
-: The study provides no evidence for masculinization of OS females/demasculinization of OS males.
+/-: The study provides evidence for masculinization of OS females/demasculinization of OS males in some cases (e.g. in some investigated measures, in some age groups or in some statistical analyses), but not in all.

Table 2 - Studies on perceptual and cognitive traits in opposite-sex and same-sex twins

Trait	**Publication**	**Setting**	**Numbers of opposite-sex (OS) and same-sex (SS) twins**	**Is there a comparison between OS and SS dizygotic (DZ) twins only?[a]**	**Age**	**Assessment**	**Findings OS vs. SS females**	**Findings OS vs. SS males**	**Suggests evidence for masculinization of OS females[b]**	**Suggests evidence for demasculinization of OS males[b]**
Academic performance	Ahrenfeldt et al., 2015a	Denmark	968 OSF 2198 SSF 966 OSM 2194 SSM (breakdown not provided)	Yes	15-16 years	Ninth grade test scores obtained from registers	OSF obtained slightly lower test scores in mathematics than SSF – opposite to what was expected	No significant differences	-	-

Auditory system functioning	McFadden, 1993	Texas, USA	17 OSF 32 SSF (DZ) 46 SSF (MZ) 17 OSM 24 SSM (DZ) 34 SSM (MZ)	Yes	OS: mean = 20.8 years SS (DZ): mean = 18.7 years SS (MZ): mean = 21.3 years	Spontaneous otoacoustic emissions (SOAEs)	SOAE frequency lower for OSF than for SSF	No significant differences	+	-
	McFadden et al., 1996	Texas, USA	18 OSF 32 SSF (DZ) 42 SSF (MZ) 17 OSM 24 SSM (DZ) 34 SSF (MZ)	Yes	OS: mean = 20.7 years SS (DZ): mean = 18.2 years SS (MZ): mean = 21.3 years	Click-evoked otoacoustic emissions (CEOAEs)	No significant differences, but spectral power of CEOAEs marginally lower (more masculinized) for OSF than for SSF	No significant differences	+	-
Expressive vocabulary	Galsworthy et al., 2000	England and Wales	974 OSF 994 SSF (DZ) 1126 SSF (MZ) 974 OSM 992 SSM (DZ) 952 SSM (MZ)	Yes	2 years	Parent report: MacArthur Communica-tive Developmental Inventory (MCDI) - Short form	OSF lower MCDI scores (more masculinized) than SSF	OSM higher MCDI scores than SSM	+	+
	Van Hulle et al., 2004	Wisconsin, USA	112 OSF 61 SSF 8DZ) 64 SSF (MZ) 112 OSM 64 SSM (DZ) 52 SSF (MZ)	Yes	20-38 months	Parent report: MCDI - Short form	OSF lower MCDI scores than SSF	OSM higher MCDI scores than SSM	+	+
Visuo-spatial cognition	Vuoksimaa et al., 2010b	Finland	120 OSF 351 SSF 110 OSM 223 SSM	Yes	Mean = 22.4 years	Mental Rotation Test (MRT)	OSF performed better than SSF	No significant differences	+	-
	Heil et al., 2011	Germany	100 OSF 100 SSF (DZ)	Yes	19-39 years, mean = 23.6 years	Mental rotation test from an MRT questionnaire	OSF performed better than SSF	NA	+	NA

NA: Not available, OS: opposite-sex, SS = same-sex, F = females, M = males, MZ = monozygotic, DZ = dizygotic, UZ = unknown zygosity

[a]Yes: The study makes a separate comparison of OS vs SSDZ twins, although it includes both MZ and DZ twins.

No: The study does not make a separate comparison of OS vs SSDZ twins, thus, including MZ twins in all analyses.

DZ twins only: The study includes DZ twins only. Thus, no MZ twins are included in the study.

[b]+: The study provides evidence for masculinization of OS females/demasculinization of OS males.

-: The study provides no evidence for masculinization of OS females/demasculinization of OS males.

+/-: The study provides evidence for masculinization of OS females/demasculinization of OS males in some cases (e.g. in some investigated measures, in some age groups or in some statistical analyses), but not in all.

Table 3 - Studies on behavioral traits in opposite-sex and same-sex twins

Trait	Publication	Setting	Numbers of opposite-sex (OS) and same-sex (SS) twins	Is there a comparison between OS and SS dizygotic (DZ) twins only?[a]	Age	Assessment	Findings OS vs. SS females	Findings OS vs. SS males	Suggests evidence for masculinization of OS females[b]	Suggests evidence for demasculinization of OS males[b]
Alcohol use	Lenz et al., 2012	A Sweden B Australia	Sample A: 29,933 females 29,271 males Sample B: 2,041 males 3,848 females (Breakdown not provided)	Yes	Not provided	Hospitalization rate due to alcohol dependence from the Swedish Twin Registry. Self-reported lifetime prevalence of alcohol dependence obtained from telephone interviews	No significant differences	Higher prevalence of alcohol dependence for OSM than SSM	-	+
	Ellingson et al., 2013	Australia	621 OSF 915 SSF (DZ) 1,191 SSF (MZ)	No	32-43 years, mean = 37.7 years	Self-reported alcohol use, alcohol dependence and abuse symptoms obtained from telephone interviews	OSF slightly more lifetime alcohol use disorder (AUD) symptoms than SSF. No other differences	NA	+/-	NA
Autistic symptoms/ hyperactivity disorder (ADHD)	Ho et al., 2005	Missouri, USA	250 OSF 142 SSF (DZ) 177 SSF (MZ) 250 OSM 128 SSM (DZ) 91 SSM (MZ)	Yes	7–15 years	Parent report: sub-threshold autistic symptomatology assessed using the Social Responsiveness Scale (SRS)	No significant differences	OSM lower mean score on the SRS (indicating less autistic symptomatology) compared with SSM	-	+
	Atterman et al., 2012	Denmark	2927 OSF 3412 SSF	Yes	3-15 years	Parental responses to items from the	OSF scored lower than SSF for both ADHD	NA	-	NA

						Child Behaviour Checklist (CBCL)	and total scores – contrary to expected			
	Eriksson et al., 2016	Sweden	4219 OSF 1808 SSF (DZ) 4219 OSM 2129 SSM (DZ)	DZ twins only	9 or 12 years	Parental telephone interview examining attention-deficit/hyper-activity disorder (ADHD) and autistic traits using Autism-Tics, AD/HD, and other Comorbidities Interventory (A-TAC)	OSF displayed fewer traits related to attention deficit and repetitive and stereotyped behaviors than SSF – opposite to what was expected	OSM displayed a larger number of traits related to attention deficit and repetitive and stereotyped behaviors than SSM	-	-
Disordered eating (DE)	Culberg et al., 2008	Michigan, USA	59 OSF 132 SSF (DZ) 172 SSF (MZ) 54 OSM 62 SSM (DZ) 103 SSM (MZ)	Yes	Mean = 20.8 years	Self-report from the Minnesota Eating Behavior Survey	Less DE for OSF compared with SSF	More DE for OSM compared with SSM	+	+
	Raevuori et al., 2008	Finland	793 OSF 765 SSF (DZ) 868 SSF (MZ) 717 OSM 705 SSM (DZ) 540 SSM (MZ)	Yes	22–28 years, mean = 24.4 years	Self-reported current and lifetime diagnoses of anorexia nervosa and bulimia nervosa obtained from Structured Clinical Interview for DSM-IV – short version by telephone	No significant differences	No significant differences	-	-
	Baker et al., 2009	Sweden	371 OSF 213 SSF (DZ) 439 SSF (MZ) 361 OSM 344 SSM (DZ) 461 SSM (MZ)	Yes	15–17 years	Self-report from Eating Disorder Inventory-2 (questionnaire)	No significant differences	No significant differences	-	-

	Lydecker et al., 2012	A: Virginia, USA B: Norway C: Sweden	Sample A: 481 OSF 408 SSF (DZ) 614 SSF (MZ) 317 OSM 295 SSM (DZ) 492 SSM (MZ) Sample B: 345 OSF 530 SSF (DZ) 900 SSF (MZ) 341 OSM 235 SSM (DZ) 445 SSM (MZ) Sample C: 2433 OSF 2901 SSF (DZ) 4099 SSF (MZ) 2423 OSM 1918 SSM (DZ) 2684 SSM (MZ)	Yes	A: Females: mean = 40.4 years Males: mean = 42.3 years B: Females: mean = 28.2 years Males: mean = 28.3 years C: Females: mean = 33.5 years Males: mean = 33.4 years	Self-reported eating disorders from interviews based on DSM-IV criteria	No significant differences when comparing OS and SS (DZ) twins but co-twin sex was associated with broadly defined bulimia nervosa in sample C when MZ twins were included	No significant differences	-	-
	Culbert et al., 2013	Michigan, USA	64 OSF 178 SSF	No	Pre-Early puberty OSF: mean = 11.7 years, SSF: mean = 11.4 years Mid-Late puberty OSF: mean = 14.0 years SSF: mean = 13.2 years	Self-reported DE attitudes assessed by the Minnesota Eating Behavior Survey and the Eating Disorder Examination Questionnaire	OSF exhibited lower levels of DE attitudes (more masculinized) during mid-late puberty compared with SSF. No significant differences during pre-early puberty	NA	+/-	NA
Feminity - masculinity	Loehlin and Martin, 2000	Australia	Sample A: 691 OSF 3158 SSF 634 OSM 1390 SSM Sample B: 447 OSF 1773 SSF 399 OSM 1041 SSM	No	A: 24–87 years, mean = 41.2 years B: 17-28 years, mean = 23.2 years	Masculinity-femininity scales “Worried”, “Reserved” and “Breaks Rules” scales, derived from the Eysenck Personality Questionnaire and Cloninger Tridimensional Personality Questionnaire	A: No significant differences B: OSF higher scores on Breaks-Rules subscale than SSF but not on the other subscales	A: OSM higher than SSM on Worried subscale and Breaks-Rules subscale but not for Reserved subscale. B: OSM higher than SSM on the Breaks-Rules subscale. No significant difference on	+/-	+/-

						(short versions)		Worried and Reserved subscale		
Feminity – masculinity/ Pubertal development/ Fertility in early adulthood	Rose et al., 2002	Finland	972 OSF 967 SSF 931 OSM 1060 SSM Sample A: 783/486 OSF 762/466 SSF Sample B: 772/464 OSF 749/454 SSF (Number of males not provided) Sample C: 4767 OSF 7528 SSF	No	16 years A: 11 years B: 14 years (follow up) C: 15-28 years	Self-report from Feminine Interest (FEM) Scale A and B: Post-menarche frequencies. B: Self-reported Pubertal Development Scale (females and males) C: Fertility in early adulthood from Central Population Records	No significant differences	No significant differences	-	-
	Verweij et al., 2016	Sweden	536 OSF 392 SSF (DZ) 695 SSF (MZ) 536 OSM 248 SSM (DZ) 374 SSM (MZ)	No	27-54 years, mean = 40.7 years	Self-reported Big-Five personality inventory data used to create a bipolar masculinity-femininity (M-F) scale	OSF scored higher (more masculine) on the M-F scale than SSF	OSM scored higher (more masculine) on the M-F scale than SSM	+	-
Masculine Attitudes (Conservatism)	Miller and Martin, 1995	Australia	905 OSF 3964 SSF 905 OSM 1832 SSM	No	Adults	Self-report from Wilson–Patterson Conservatism Scale (questionnaire)	OSF reported more masculine (socially conservative) views compared with SSF	No significant differences	+	-
Religiousness and religious coping	Ahrenfeldt et al., 2016	Denmark	408 OSF 1383 SSF 350 OSM 856 SSM (Number of MZ twins not provided)	Yes – comparison between OS and SS (DZ)+UZ twins	20-40 years, mean = 29.8 years	Self-reported survey including questions on religiousness and religious coping	No significant differences	No significant differences except that OSM reported attending church more often than SSM in childhood	-	-

Sensation seeking/ agression	Resnick et al., 1993	England	51 OSF 286 SSF 51 OSM 85 SSM	No (scores for MZ and SSDZ twins did not differ)	16-70 years	Self-report from Sensation Seeking Scale IV (SSS-IV) – an overall score and four subscales scores (questionnaire)	OSF reported increased disinhibition, experience seeking and overall sensation seeking compared with SSF	No significant differences	+/-	-
	Cohen-Bendahan et al., 2005a	The Netherlands	74 OSF 55 SSF	DZ twins only	13 years	Self-report from Sensation Seeking Scale V (SSS-V) questionnaire and from Olweus Multifaceted Aggression Inventory (OMAI) Test of child: Reinisch Aggression Inventory (RAI)	OSF lower experience seeking behavior – opposite to expected. OSF rated more aggressive than SSF on two of the RAI but no differences on OMAI subscales	NA	-/+	NA
	Slutske et al., 2011	Australia	564 OSF 836 SSF (DZ) 1,111 SSF (MZ)	Yes	32-43 years, mean = 37.7 years	Telephone interview and questionnaire: The Zuckerman Sensation Seeking Scale-Form V (SSS-V)	OSF showed higher experience-seeking and thrill-and – adventure seeking scores compared with SSF. Other group differences were non-significant	NA	+/-	NA
Sex-typed play	Elizabeth and Green, 1984	New York and Long Island area	156 OSF 270 SSF (DZ) 320 SSF (MZ) 156 OSM 240 SSM (DZ) 262 SSM (MZ)	Yes	4-12 years, mean = 6.8 years	Parental report: a 17-item questionnaire used to discriminate between typical and atypical sex-role development	No significant differences	OSM scored in a more feminine direction compared with SSM	-	+

	Henderson and Berenbaum, 1997	Illinois, USA	35 OSF 36 SSF	DZ twins only	3–8 years, OSF: mean = 5.1 years SSF: mean = 5.5 years	Observational: time child played with toys typically preferred by boys/girls. Mothers' reports of game preferences	No significant differences	NA	-	NA
	Rodgers et al., 1998	Oregon, USA	16 OSF 54 SSF 16 OSM 48 SSM	DZ twins only	7–12 years, OS: mean = 8.3 years SSF: mean = 8.1 years SSM: mean = 8.5 years	Observational: toy preferences during 5-minute free-play interactions between mothers and twins	No significant differences	No significant differences	-	-
Social skills	Laffey-Ardley and Thorpe, 2006	Australia	36 OSF 28 SSF (DZ) 36 OSM 22 SSM (DZ)	DZ twins only	3–6 years, OS: mean = 4.4 years SS: mean = 4.7 years	Parent and teacher reports of social co-operation and competence using the Preschool and Kindergarten Behavior Scale — 2nd Edition (PKBS-2)	Parent reports: Mixed result for social co-operation: OSF rated higher than SSF when using adjusted ANCOVA analysis but not for simple ANOVA analysis	Teacher-reports: SSM rated higher than OSM on social co-operation	+/-	-

NA: Not available, OS: opposite-sex, SS = same-sex, F = females, M = males, MZ = monozygotic, DZ = dizygotic, UZ = unknown zygosity

[a]Yes: The study makes a separate comparison of OS vs SSDZ twins, although it includes both MZ and DZ twins.

No: The study does not make a separate comparison of OS vs SSDZ twins, thus, including MZ twins in all analyses.

DZ twins only: The study includes DZ twins only. Thus, no MZ twins are included in the study.

[b]+: The study provides evidence for masculinization of OS females/demasculinization of OS males.

-: The study provides no evidence for masculinization of OS females/demasculinization of OS males.

+/-: The study provides evidence for masculinization of OS females/demasculinization of OS males in some cases (e.g. in some investigated measures, in some age groups or in some statistical analyses), but not in all.

Table 4 - Studies on other health traits in opposite-sex and same-sex twins

Trait	Publication	Setting	Numbers of opposite-sex (OS) and	Is there a comparison between OS	Age	Assessment	Findings OS vs. SS females	Findings OS vs. SS males	Suggests evidence for masculiniza-	Suggests evidence for demasculiniza-

			same-sex (SS) twins	and SS dizygotic (DZ) twins only?[a]					tion of OS females[b]	tion of OS males[b]
Cancer	Ahrenfeldt et al., 2015b	Denmark and Sweden	44,650 OSF 84,721 SSF 44,660 OSM 88,261 SSM	Yes – comparison between OS and SS (DZ)+UZ twins	0-73 years Denmark: mean = 25 years at start of follow up Sweden: mean = 24 years at start of follow up	All-cause cancer and sex-specific cancers identified from the Danish and Swedish twin and cancer registries	No significant differences	No significant differences	-	-
Early life mortality risks	Ahrenfeldt et al., 2017	Denmark	12,033 OSF 12,051 OSM 21,644 SSF 22,901 SSM	No	0-15 years	Perinatal mortality, neonatal and postneonatal deaths and child mortality identified through the Danish Twin Registry	OSF lower perinatal and early neonatal mortality than SSF	OSM lower perinatal, neonatal and postneonatal mortality than SSM - may be due to MZ twins in the SS twin group	-	+/-
Epilepsi	Mao et al., 2018	Denmark	11,078 OSF 19,186 SSF 11,080 OSM 20,207 SSM	Yes	0-34 years	Epilepsy identified through the Danish National Patient Registry	No significant differences	No significant differences	-	-

NA: Not available, OS: opposite-sex, SS = same-sex, F = females, M = males, MZ = monozygotic, DZ = dizygotic, UZ = unknown zygosity

[a]Yes: The study makes a separate comparison of OS vs SSDZ twins, although it includes both MZ and DZ twins.

No: The study does not make a separate comparison of OS vs SSDZ twins, thus, including MZ twins in all analyses.

DZ twins only: The study includes DZ twins only. Thus, no MZ twins are included in the study.

[b]+: The study provides evidence for masculinization of OS females/demasculinization of OS males.

-: The study provides no evidence for masculinization of OS females/demasculinization of OS males.

+/-: The study provides evidence for masculinization of OS females/demasculinization of OS males in some cases (e.g. in some investigated measures, in some age groups or in some statistical analyses), but not in all.