

The beauty, logic and limitations of twin studies

Espen Røysamb^{1,2} and Kristian Tambs²

1) Department of Psychology, University of Oslo, Norway

2) Norwegian Institute of Public Health, Norway

Correspondence: Espen Røysamb, Department of Psychology, University of Oslo, P.O. box 1094 Blindern, NO-0317 Oslo, Norway
email: espen.roysamb@psykologi.uio.no

ABSTRACT

During the last fifty years more than 2700 twin studies have been published, examining the etiology of a high number of traits. Twin studies enable investigation of both genetic and environmental effects, and thereby also examination of causal factors involved in human traits and disorders. The beauty of twin studies resides in the potential of studying the unobserved by the logic of a design. The aim of this article is to outline central theoretical foundations and possible limitations, and to review selected key findings. We describe the inherent fundamentals of the classic and extended twin designs. The logic of the main analytic approaches is outlined, and the principles of univariate biometric modelling described. Next, we review different multivariate models, including the Cholesky, correlated factors, common factor, common pathway and phenotypic causality models. Additionally, the cotwin-control approach, representing a natural experimental design, and mimicking a counterfactual situation, is outlined. Central assumptions, threats and limitations of the twin design are discussed. In particular, we address the issue of missing heritability, non-random mating, the equal environment assumption and gene-environment correlations. Finally, we review some selected findings from the field of behavior genetics and twin studies.

This is an open access article distributed under the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Why do people differ on physical and psychological traits? How do nature and nurture contribute to human variation in physical height, cardiovascular risk, personality, depression or intelligence? These are fundamental questions raised by lay people, philosophers and scientists alike. Progress in science has led to methodological designs and subsequent empirical findings that answer some of these questions. In this article we aim to outline the principles of the twin design, as a method for studying genetic and environmental influences on human traits, and review some exciting findings and implications. The beauty of twin studies lies in the potential to examine unobserved causes by the logic of a design.

Twin studies have become a major source of knowledge about heritability for various traits and conditions (Boomsma et al, 2002; Plomin et al, 2013). A recent meta-analysis of the heritability of human traits, published in *Nature Genetics*, revealed that during the last fifty years a total of 2,748 publications have reported twin correlations and variance components for 17,804 traits, based on 14,558,903 partly independent twin pairs (Polderman et al, 2015). These numbers testify to the importance and value of twin studies during a half century of remarkable scientific progress.

How can we study unobserved entities and phenomena? At a general level, investigating the unobserved has been a subject for astronomers studying black holes, physicists studying quantum reactions, or physicians studying diseases for which only the symptoms could be identified. This has also been one of the challenges facing researchers interested in delineating genetic and

environmental influences on human traits (Plomin et al, 2013). The twin design represents a way of studying effects without directly observing genes and environments.

Another basic challenge for many scientific endeavors is to establish evidence of causality. Within psychology, medicine and epidemiology the experimental design, preferably randomized and double blinded, represents the ideal approach to causal evidence (Plomin et al, 2013; Rutter, 2007). However, for many research questions the randomized experiment is neither ethical nor feasible. The twin design represents an approach in which effects of causal factors may be estimated from cross-sectional data. That is, twin analyses disentangle the genetic and environmental factors that contribute to the variance in a certain phenotype, and as such are causally related to the phenomenon in focus.

Twin studies are often seen as primarily informative of genetic effects. However, a noteworthy feature of the twin designs is the capacity to establish evidence of the causal role of environmental factors. Many epidemiological, psychological and psychiatric studies aim to study causal environmental effects without being designed to do so. For example, studies have shown associations between 'exposures' such as education or life events and mental health outcomes (Bannink et al, 2013; Dalgard et al, 1995; Flouri & Kallis, 2011; Kinderman et al, 2013; Mezuk et al, 2008), and such associations often are interpreted as evidence of causal effects. However, as there are substantial genetic influences on both educational level and certain types of life events (Heath et al, 1985; Kendler & Baker, 2007; Kendler et al, 1999; Tambs et al, 2012) the associations may be partly or entirely explained by genetic

confounding. Twin designs allow controlling for genetic factors and thereby identification of effects that are truly environmental in nature.

The aim of this paper is to a) outline the basic principles and logic of the twin designs, b) comment on limitations and critical aspects, and finally c) review selected key findings.

CLASSIC AND EXTENDED TWIN DESIGNS

Twins constitute about three percent of the population (Martin et al, 2012), and include both identical (monozygotic, MZ) twins and fraternal (dizygotic, DZ) twins. The MZ/DZ ratio is roughly 1:2, but there is some variation across nations and ethnic groups, and the proportion of DZ twins increases with in vitro fertilization (D'Addato, 2007; Imaizumi, 2003; Tandberg et al, 2007). MZ twins share 100% of their genes, while DZ twins, just like ordinary siblings, share, on average, 50% of their segregating genes (Plomin et al, 2013). Accordingly, MZ cotwins are always the same sex, whereas DZ twin pairs can be either same-sex or opposite sex.

The *classic twin design* (Boomsma et al, 2002) includes MZ and DZ twins reared together, and the within-pair similarities represent the cornerstone of subsequent analyses. That is, we examine and compare the degree of twin-cotwin similarity among MZ pairs versus DZ pairs. Similarity is typically operationalized as correlations (pearson, polychoric or tetrachoric) and sometimes as concordance rates. Based on the observed correlations we estimate the contributions of genetic and environmental factors. More specifically, we test alternative theoretical models against the observed data to examine the presence and magnitude of effects (Neale & Cardon, 1992; Plomin et al, 2013).

The classic twin design may be *extended* in several ways. First, one option includes *twins reared apart*, that is, one or both twins in a pair have been adopted. Studies with twins reared apart yield additional evidence of genetic versus familial environmental sources of twin similarity. Second, twin samples may be supplemented with *other types of relatives*, such as siblings, parent-offspring and potentially other relatives. Full siblings share 50% of segregating genes, as DZ twins do, and comparison of these groups allows examination of twin-specific factors. When including relatives such as siblings of MZ twins, or parents of twins, we are able to distinguish between additive and non-additive genetic effects while also estimating common environmental factors. Third, *longitudinal twin designs* enable examination of stability, development and change. The genetic and environmental influence on the stable variance in a phenotype may be estimated, and correspondingly influences on change and development. Interestingly, several studies have reported high genetic influences on stability factors, but also genetic innovation – new genetic factors starting to operate at certain ages (Kendler et al, 2008; Nes et al, 2006). Fourth, the *children of twins design* involves

offspring of MZ and DZ twins, and examines similarity across and within generations. Since MZ cotwins have identical genes, children of MZ twins are on average as strongly genetically related with their MZ aunts as with their mothers, or with their MZ uncles as with their fathers. A child of an MZ twin mother may resemble her mother and maternal aunt equally much, thereby suggesting purely genetic sources of the resemblance, or may resemble her aunt less than her mother, thus suggesting environmentally driven resemblance. Fifth, twin studies may be combined with *molecular genetic* research, in order to both obtain estimates of total genetic effects and delineation of specific genetic influences (Boomsma et al, 2002; De Neve et al, 2012; Rietveld et al, 2013).

Although the extended twin designs provide additional power and opportunities compared to the classical twin design, the latter is thus far more common. However, the extended designs offer an important validation of findings from the classic designs, for example when findings from classical twin designs are replicated and confirmed in twin-adoption studies.

THE LOGIC OF THE MAIN ANALYTIC APPROACHES

Biometric modelling – univariate models

In theory, four general types of factors may affect human traits: Additive genetic effects (A), non-additive genetic effects (D: dominance/epistasis), shared (or common) environmental effects (C), and non-shared environmental effects (E) (Plomin et al, 2013). Figure 1 illustrates how these factors influence a certain phenotype in a pair of twins. These four factors all have a unique profile of associations within MZ versus DZ twin pairs. The A effects are correlated at unity among MZ twins and at $r=0.5$ among DZ twins (due to the fact that MZ twins share 100% of genes, and DZ twins share on average 50% of segregating genes). The D effects include interaction at specific loci (dominance) or across loci (epistasis) and are also perfectly correlated among MZ twins and at 0.25 among DZ twins (if D only includes dominance, and typically somewhat lower than 0.25 in the presence of multi-loci epistasis). The C effects are per definition the environmental factors that contribute to similarity among both MZ and DZ twins, and are thus correlated at unity for both zygosity groups. Finally, the E effects represent all components of experience and stimuli that contribute to variance but not covariance, that is, contribute to dissimilarity between both MZ and DZ twins. Hence, E effects are uncorrelated for both zygosity groups. Note that the boundaries between what contributes to effects from C and E are usually rather subtle. C effects do not necessarily include all environmental factors or events to which both twins in a pair are exposed. Rather, C includes factors that in consequence contribute to similarity, and are shared in effect rather than in observable exposure. For instance sharing school class

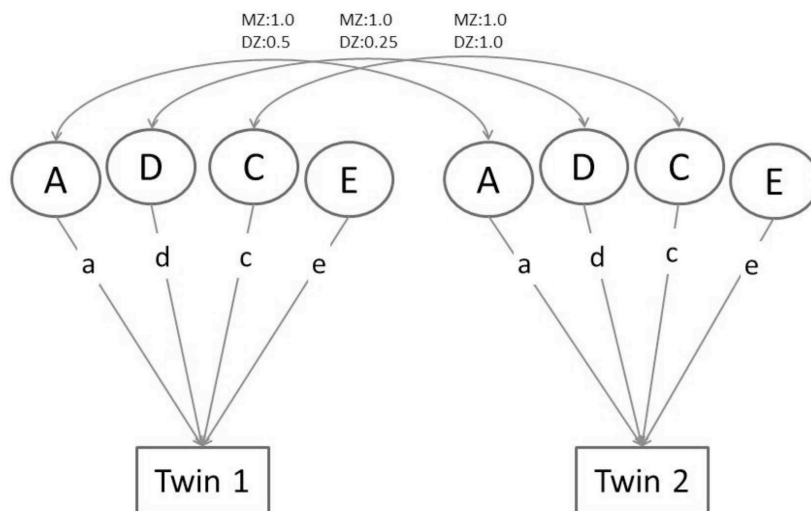


Figure 1. The classic univariate twin model. A=Additive genetic effects; D=Dominance/epistasis; C=Common/shared environment; E=Non-shared environment; MZ=Monozygotic; DZ=Dizygotic.

may be thought of as contributing only to C, but since not all experiences in a class are equal for all pupils, environmental effects associated with going to a certain school class can contribute to both C and E effects.

Figure 1 is drawn according to the rules of path analysis, a method used extensively in social sciences today, but developed for the use in quantitative genetics almost a 100 years ago (Wright, 1921, 1934). The figure shows the theoretical correlation structure of the latent A, D, C and E factors for any given phenotype. The path coefficients *a*, *d*, *c* and *e* are free parameters and can be thought of as regression coefficients. They represent the effect of the latent factors on the phenotype, whereas the inter-factor correlations are fixed according to known associations for MZ and DZ twins (e.g. for A-factors the correlations are 1.0 and 0.5 for MZ and DZ pairs respectively).

Based on the theoretical expectations, and according to the rules of path analysis, the phenotypic variances (*V*), standardized to unity, and the correlations (*r*) between cotwins for a certain phenotype, can be expressed as

$$\begin{aligned} \text{I } V_{DZ} &= V_{MZ} = a^2 + d^2 + c^2 + e^2 = 1 \\ \text{II } r_{MZ} &= a^2 + d^2 + c^2 \\ \text{III } r_{DZ} &= 0.5 a^2 + 0.25 d^2 + c^2 \end{aligned}$$

where the latent factors A, D, C and E are all standardized, and *a*, *d*, *c* and *e* are the (standardized) path coefficients shown in Figure 1. The three equations above include four unknown parameters, and in order to solve them, we need to assume that one parameter is zero. Combining equations II and III can help us decide which parameter to drop and which to keep. The expression (2* Equation III – Equation II) gives

$$2r_{DZ} - r_{MZ} = c^2 - 0.5 d^2$$

Since neither *c*² or *d*² can take on negative values, we keep *c* and drop *d* if $2r_{DZ} - r_{MZ} > 0$, that is, if *r*_{MZ} is smaller than twice the *r*_{DZ}. We keep *d* and drop *c* if $2r_{DZ} - r_{MZ} < 0$, that is, if *r*_{MZ} is larger than twice the *r*_{DZ}.

Further, the equations may be solved for the different parameters, and we obtain the following (assuming no *d* effect):

$$\begin{aligned} a^2 &= 2(r_{MZ} - r_{DZ}) \\ c^2 &= 2r_{DZ} - r_{MZ} \\ e^2 &= 1 - r_{MZ} \end{aligned}$$

Traditionally the genetic variance component, the heritability, is denoted *H*² as a sum of additive, *a*², and non-additive, *d*², genetic effects (‘broad heritability’), or *h*² in models without a *d* (‘narrow heritability’). In wording, the expressions above state that the heritability, *a*² or *h*², equals twice the difference in MZ vs DZ correlations (i.e. ‘narrow heritability’) and the effects of the shared and non-shared environmental factors are also given by these simple equations (Plomin et al, 2013). Note that whereas the formulas yield simple estimates of the various effects, no measure of uncertainty (standard errors or confidence intervals) is provided by the equations as such.

These equations and the model depicted in Figure 1 represent the foundation for empirical model testing. By means of structural equation modelling (SEM) we fit theoretical models to the observed data to obtain parameter estimates, confidence intervals and model fit. SEM software such as Mplus, Lisrel or Amos may be used to test such models (Hoyle, 2012; Loehlin, 2004). However, programs have been developed with special features advantageous for twin designs, in particular the freely available Mx software, and more recently the open source package OpenMx that operate within the R framework (Boker et al, 2011).

As already described the ACDE-model, as depicted in Figure 1, is over-parameterized in the classical twin design, meaning that only reduced models, such as an ACE or ADE model, will have sufficient degrees of freedom to permit parameter estimations. However, the observed correlation structure (e.g. DZ correlation higher or lower than half the MZ correlation) is highly informative as to which model is plausible. Further,

with extended designs, including for example twins reared together and apart, full ACDE models may be tested.

Sex-differences in genetic and environmental effects may also be examined (Neale et al, 2006; Plomin et al, 2013). *Quantitative* sex-differences, referred to as common sex-limitation, imply that the same genetic and environmental factors operate for women and men, but that the magnitude of effects may vary across sex. *Qualitative* sex-differences, or general sex-limitation, imply not only magnitude differences, but also that partly different genetic, or environmental, factors operate for women and men. Quantitative differences are indicated if the MZ and DZ correlations, and the ratio between them, differ across sexes, whereas qualitative differences are suggested if the twin-cotwin correlation among opposite sex DZ twins is lower than among same sex DZ twins. Within the major types of quantitative and qualitative differences there are subtypes including scalar models in which the standardized variance components may be equal across sex, but the total variance and unstandardized components may differ. The sex-limitation models can address important questions about the etiology of phenotypes across sex, yet, these models cannot address questions of prevalence differences between sexes. For example, twin models may tell us to what extent the same underlying factors – genetic or environmental – influence depression in women and men, and also if the heritability is the same, but do not provide direct evidence pertaining to sex-differences in prevalence or level of depression.

Multivariate biometric models

The basic univariate model, as described above, may be extended into various multivariate models. Generally, whereas the univariate model is set up to estimate sources of variance in a given phenotype, the multivariate models also provide estimates of the sources of *covariation* between phenotypes (Plomin et al, 2013). For example, research has shown substantial correlations, or comorbidity, between disorders such as anxiety and depression. A multivariate twin study can delineate the genetic and environmental contribution to this association. In the case of anxiety and depression, studies have revealed that genetic factors represent a major source of covariation, or a shared liability for both anxiety and depression, and that environmental factors are primarily important for the uniqueness of the two conditions (Kendler et al, 1992, 2007).

Several variants of the multivariate model have been developed, each with features designed to address specific research questions. Figures 2a-e show commonly used models. For ease of communication we present the bivariate case where we show the manifest and latent variables for one twin only, and only include A and E factors. Yet, the model could in principle include a high number of phenotypes, and all the A, D, C and E factors.

All models include two observed variables (phenotypes) X and Y, and reflect different ways of conceptualizing the structure of genetic and environmental factors, represented as latent A and E variables. Figure 2a shows a *Cholesky model* (Carey, 2004; Loehlin, 1996) in which one set of latent variables (A and E) affect X, and also influence Y, whereas another set is unique to Y. The Cholesky model represents an efficient way of decomposing the multivariate correlation structure, and can be advantageous for handling problems of model identification and non-positive definiteness. However, there are certain problems with this approach when testing sex-limited models (Neale et al, 2006), and often result presentations would benefit from translation into another model (Loehlin, 1996). Figure 2b shows a *correlated factor model*, in which the associations between latent factors are modelled explicitly as factor correlations. In Figure 2c an *independent factor model* is shown, where one set of latent variables, A and E, is shared by the phenotypes, and another set is unique to each. Next, Figure 2d shows *the common factor model*, where the two observed variables share an underlying phenotypic factor (F) which again is influenced by A and E factors. Finally, Figure 2e shows the *phenotypic causation model* with separate A and E factors for the two phenotypes, and the association between them accounted for by a single causal path. This model is theoretically important due to its potential to address questions of causality, yet there are some challenges of identification and interpretation of the model. Note also that models a, b and c are different variants of the same structure, that is, they may all estimate heritabilities, genetic and environmental correlations, and will fit to the same degree in the bivariate case. Either model may be translated into one of the other (Loehlin, 1996). In contrast, model d) tests the hypothesis that shared A and E effects are mediated through a common phenotypic factor, and model e) tests a hypothesis of phenotypic causation.

Cotwin control design

A related, yet conceptually somewhat different approach to twin analysis is represented by the cotwin-control method (McGue et al, 2010; Rutter, 2007). This approach is based on the *counterfactual* model, where we basically ask if outcome Y would have occurred if exposure X had (not) happened. As a thought experiment we might wonder if a certain person would have developed a depressive disorder if she had not lost her job and gone through a difficult divorce. Yet, a person cannot both have been exposed and not been exposed to the same event. An experimental design could in theory inform us about the causal effects of factors such as unemployment and divorce, but such experiments are not feasible. However, two MZ twins are genetically identical and share the same family background, yet may be discordant in certain exposures (e.g. divorce) and as such may re-

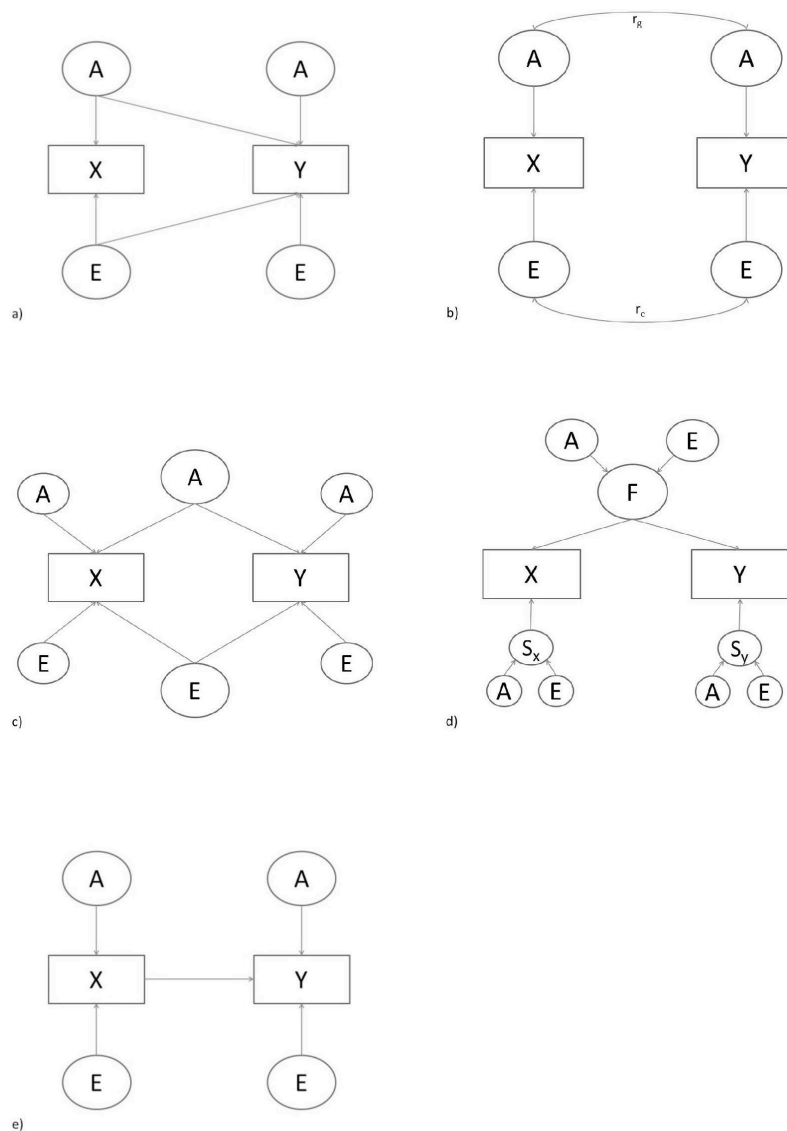


Figure 2. Bivariate twin models: Cholesky model (2a), correlated factor model (2b), independent factor model (2c), common factor model (2d), and phenotypic causality model (2e).

present a counterfactual scenario (McGue et al, 2010; Sadler et al, 2009).

Generally, for deduction of causality, standard observational approaches are somewhat limited due to issues of confounding, reverse causation and possible nonrandom distribution of exposures. Natural experiments represent a broad class of designs that might be helpful when randomized blinded experiments are not feasible, and the cotwin control or discordant-twin-pair design constitutes one such approach (McGue et al, 2010; Rutter, 2007). In this design we compare the individual level association between exposure and outcome with the corresponding within-MZ (and DZ) association. If an exposure X is associated with outcome Y by the magnitude B_1 when analyzed at the individual level – as in any standard regression model with an ordinary sample (or in this case a twin sample but with disregard of the twin structure), and the X-Y association remains equally strong when analyzed within MZ pairs (i.e., $B_1=B_{MZ}$) the effect is seen as due

to non-shared environmental mechanisms (in classical twin studies termed E). For example, if smoking (X) is found to statistically predict a respiratory disease (Y) in the population, and we find a similar effect size among MZ-twins discordant for smoking, this finding fits with a model of smoking being causally related to the disease. If, on the other hand, we observe that the initial association B_1 vanishes in the within pair analyses (e.g. the smoking twin has no higher risk of subsequent disease Y) the initial association is probably due to confounding by genetic factors or family environment.

The scenarios described are extreme versions in the sense of referring to confounding as present or absent, but we could observe any degree of genetic/familial confounding. Also, although discordant MZ twins are crucial to the analyses the DZ twins are important in confirming a pattern of confounding or causality. The cotwin control approach is elegant in its way of mimicking a counterfactual situation and thereby an

experimental design. However, in its basic form the design focuses on discordant twin pairs only, thus disregarding concordant pairs and reducing statistical power. Newer approaches exploit all degrees of twin differences, and may examine discordance in either exposure or outcome (De Moor et al, 2008; McGue et al, 2010). Compared to the biometric model described above the cotwin control design is not a totally different approach. Rather, the potentially causal effect of X on Y would appear in a bivariate biometric analysis as well, for example as the cross-effect of E on Y in Figure 2a. As in the biometric models, the twin structure provides unique opportunities for examining genetic and environmental influences, and although final evidence of causality is not obtained, results may support or reject a causal model.

LIMITATIONS OF THE TWIN DESIGNS

Missing heritability

The major advantage of the twin designs is the ability to capture and estimate the total genetic and environmental effects on a certain phenotype and on the covariation between several phenotypes. Designed initially to estimate variance-covariance components, twin studies per se do not identify specific genes. Correspondingly, although the total effect of shared and non-shared environment is estimated, specific environmental factors that contribute to this variance are rarely identified, although specific effects of observed environmental factors, like for instance life events, can be built into the model. Thus, an obvious limitation of most twin studies is the lack of identification of specific genes or environments involved in a certain phenotype.

Other approaches, such as Genome Wide Association Studies (GWAS), are designed to search for specific genes (Donnelly, 2008; Pearson & Manolio, 2008). So far, many initial findings have not been replicated, and current evidence point in a direction of a high number of Single Nucleotide Polymorphisms (SNPs) involved in even simple phenotypes. Using body height as an example, a recent study including 253,288 individuals, identified 697 genetic variants clustered at 423 loci (Wood et al, 2014). However, in validation samples the sum of effects from all the best SNPs accounted for 21-29% of variance in height, and this figure stands somewhat in contrast to the heritability of height at 80-90% when estimated in twin studies (Silventoinen et al, 2003; Wood et al, 2014; Yang et al, 2010) as well as extended twin studies (Tambs et al, 1992). The discrepancy between total effects from identified genes and the twin study heritabilities has been denoted *missing heritability* (Goldman, 2014; Turkheimer, 2012).

Do the findings of limited variance explained by identified genes, and thereby substantial missing heritability, represent a threat to the general findings of twin studies, or do they reflect current limitations of molecular genetics? In general there is reasonable

agreement that the missing heritability phenomenon is not due to invalid findings from twin studies, but rather that current techniques to examine effects of single genes are limited. GWA studies typically only capture common variants of SNPs, do not capture interactions between loci, and require extremely large sample sizes to obtain sufficient power to identify small effects when using necessary GWA-controlled significance levels (Goldman, 2014; Plomin, 2013; Reynolds & Finkel, 2015; Yang et al, 2010). We will argue that the current missing heritability should also be seen as parallel to *missing environmentality*. That is, twin studies might report 60% of the phenotypic variance in depression or wellbeing as due to environmental factors (Kendler et al, 2011; Kendler et al, 1987; Nes & Røysamb, 2015), but there is limited evidence of specific environmental exposures that causally and independently explain this variance. Thus, as the quest for specific genes and environments continues twin studies are unique in the capacity to capture, and control for, the entire genetic effect, and correspondingly the entire environmental effect.

Non-random mating

People tend to fall in love, marry, and have children with people that resemble themselves (Heath et al, 1985; Merikangas, 1982). This phenomenon, known as non-random or assortative mating, pertains not only to physical characteristics such as height and weight, but has been shown for mental disorders such as depression, anxiety, and substance abuse (Merikangas, 1982; Nordsletten et al, 2016), personality traits and life satisfaction (Ask et al, 2013) and health behaviors such as smoking, drinking and physical exercise (Ask et al, 2012). While highly interesting as a phenomenon per se, non-random mating also has potentially important implications for twin studies.

A basic assumption in twin modelling is the biological relatedness of 100% and 50% of MZ and DZ twins, respectively. What happens with this assumption if parents resemble each other and non-random mating leads to genetic similarity in parents? The genetic similarity of 50% for DZ twins is based on a notion of random mating – parents do not share genes beyond that expected by random chance. If non-random mating results in genetic similarity between parents (e.g. she and he share a genetic disposition to extraversion), then DZ twins would share more than 50% of the genes for traits that are subjected to such mating mechanisms. Note that MZ twins will share 100% of their genes, regardless of parental similarity.

If the assumption of random mating is violated, estimates from biometric twin modelling might be biased in the direction of over-estimated effects from shared environment, and under-estimated genetic effects. For example, if a certain trait has a true heritability of 50% and no effect from shared environment, under the random mating assumption we would expect to observe an MZ-correlation of .50 and a DZ-correlation of .25. Now, if our empirical data show correla-

tions of .50 (MZ) and .30 (DZ), we would (incorrectly) estimate the heritability to be 40% (i.e. $h^2 = 2(.50 - .30) = .40$) and the effect of shared environment to be 10% ($c^2 = 2 \times .50 - .50 = .10$). Thus, if assumptions of genetically random mating are violated, heritability may be under-estimated rather than over-estimated. However, biometric modelling can incorporate non-random mating effects, either in an extended twin design including data on twins and their parents, or by modelling mating effects based on estimates from other studies (Heath et al, 1985; Plomin et al, 2013). In sum, non-random mating might lead to somewhat biased (reduced) estimates of genetic effects, but may be taken into account in biometric modelling strategies.

Equal environment assumption (EEA)

Another general supposition of twin modelling is known as the equal environment assumption (Plomin et al, 2013; Richardson & Norgate, 2005). We assume that MZ and DZ co-twins are exposed to common environmental factors to the same degree. If MZ cotwins are treated similarly, for example being dressed alike, and more so than DZ cotwins, the EEA may be violated. Likewise, if MZ cotwins spend more time together or feel closer than DZ cotwins, and this stronger closeness results in higher mutual influence, often termed “cotwin cooperation” in MZs, the EEA may also be violated. In such cases excess MZ versus DZ twin similarity estimated as genetic effects, may in fact partly be due to similarity in parental treatment or cotwin cooperation. Thus, the EEA represents a basic premise of the twin model, and violations may lead to biased estimates of genetic and environmental effects.

Yet, EEA-violations can only be a threat to the validity of twin studies if 1) more of the environment is actually shared by MZ than by DZ cotwins, and 2) this type of shared environment causally influences the phenotype in focus. For example, if MZ twin children are dressed more alike than DZ twins, but dressing is uncorrelated with the study phenotype (e.g. depression, personality or education level), then the EEA is still valid.

The EEA may be examined by measuring parental treatment of twins, and the extent of contact and closeness among twins. Empirical analyses thus far provide evidence that the EEA generally holds (Derks et al, 2006; Kendler, 1993; Kendler et al, 1993; Plomin et al, 2013; Tambs et al, 1995) although there is some evidence pointing in the opposite direction for specific traits (Rose & Kaprio, 1988; Rose et al, 1990). Also, support for the validity of findings is implied by consistent results from classic twin studies and studies of twins reared apart. The majority of studies have used samples of twins reared together, but evidence from twin- and sibling-adoption studies in general support findings from the former (Lykken & Tellegen, 1996; Matteson et al, 2013; Tellegen et al, 1988).

The notion of *evocative gene-environment correlation* is also important for the EEA (McGue et al,

2006). Environmental responses may be activated by genetic dispositions in individuals. People who are friendly and smiling activate other responses than antagonistic and aggressive people. Therefore, if MZ twins, who have identical genetic dispositions, tend to evoke similar social responses, then the similarity in environmental exposure is genetic in origin. Thus, MZ twins may be treated similarly because their genetic similarity evokes similar responses in others. The mechanisms of evocative gene-environment correlation is important for understanding etiological processes, but such a correlation does not represent a violation of the EEA.

We have reviewed central assumptions of the twin design, and some limitations and critical aspects. Other related topics include gene-environment interaction, negative confounding between C and D factors, low power to detect C effects, heritability estimates from Genome-wide Complex Trait Analysis (GCTA), and finally epigenetic mechanisms. It is beyond the scope of this paper to address all these issues, but we refer the interested reader to some relevant literature (Boomsma et al, 2002; Canli & Lesch, 2007; Krishna Kumar et al, 2016; Moffitt et al, 2006; Petronis, 2010; Plomin et al, 2013; Tsankova et al, 2007; Yang et al, 2013).

SELECTED FINDINGS

What is the current state of knowledge – what do we know about human nature that we did not know a few decades ago? The recent meta-analysis by Polderman et al (2015) in *Nature Genetics* provides some important findings. Genetic influences have been established for a wide range of phenotypes, including blood pressure, metabolic functions, immunological system functions, height and weight, personality traits and disorders, conduct disorder, depressive episodes and alcohol use. Heritability estimates vary across traits, and the mean heritability is 49%. Further, for most of the traits examined (69%) the observed twin correlations are consistent with a simple and parsimonious model in which twin resemblance is accounted for by additive genetic factors (Polderman et al, 2015).

Twin studies are published in specialty journals such as *Twin Research and Human Genetics*, and *Behavior Genetics*, but increasingly also in general and top ranked journals such as *JAMA Psychiatry*, *American Journal of Psychiatry*, *Journal of Personality and Social Psychology*, *Journal of Abnormal Psychology*, and *Psychological Medicine*. As the audiences of twin studies have grown, major findings have been integrated into theoretical models of human development and disorders (Plomin et al, 2013; Rutter, 2007; Rutter et al, 2006). In addition to the general aim of delineating genetic and environmental factors, twin samples may be used as ordinary population samples – albeit with the requirement of controlling for lack of independence between observations (Røysamb et al, 2011).

Within psychology and medicine there has recently been a replication crisis (Maxwell et al, 2015; Schooler, 2014; Stroebe & Strack, 2014). Prior findings have failed to replicate in new studies, and questions have been raised as to the validity of single-study findings. With this backdrop to the current state of scientific affairs, it is noteworthy that several basic findings from twin studies have been replicated repeatedly. A recent study by Plomin and colleagues (2016) outlined *the top 10 replicated findings from behavioral genetics*. The findings can be summarized as follows:

1. All psychological traits show significant and substantial genetic influence
2. No traits are 100% heritable
3. Heritability is caused by many genes of small effect
4. Phenotypic correlations between psychological traits show significant and substantial genetic mediation
5. The heritability of intelligence increases throughout development
6. Age-to-age stability is mainly due to genetics
7. Most measures of the “environment” show significant genetic influence
8. Most associations between environmental measures and psychological traits are significantly mediated genetically
9. Most environmental effects are not shared by children growing up in the same family
10. Abnormal is normal

See Plomin et al. (2016) for an elaboration of each finding. Generally, these results represent important scientific knowledge about human lives, and testify as such to the value of behavior genetics and twin studies.

Norway has, in line with the other Nordic countries, developed a rather strong line of twin research. Based on general population registries and the establishment of a national twin registry, a number of twin projects have been developed (Harris et al, 2006; Nilsen et al, 2013; Tambs et al, 2009). We did a search in the Web of Science database, and identified more than 200 twin

studies conducted by Norwegian researchers, with more than 5400 citations. Also, the recent meta-analysis by Polderman et al (2015) identified Norwegian researchers among the top international contributors (i.e. 25+ papers) to the field of twin studies across the last fifty years.

Norwegian studies have examined genetic and environmental influences on *anxiety and depression* (Ask et al, 2016; Czajkowski et al, 2011; Nes et al, 2007; Reichborn-Kjennerud et al, 2002; Tambs et al, 2009), *personality disorders* (Gjerde et al, 2011; Kendler et al, 2008; Ørstavik et al, 2012; Reichborn-Kjennerud et al, 2007, 2010, 2015; Torgersen et al, 2008), *wellbeing and happiness* (Nes & Røysamb, 2015; Nes et al, 2006; Røysamb et al, 2002, 2003, 2014), *educational attainment* (Heath et al, 1985; Ørstavik et al, 2014; Tambs et al, 2012), *intelligence* (Sundet et al, 1994, 2005), *pain* (Nielsen et al, 2008; Reichborn-Kjennerud et al, 2002; Vassend et al, 2011), *personality traits* (Vassend et al, 2013), *sleep problems* (Nes et al, 2005) *body mass index* (Harris et al, 1995), *smoking, alcohol and drug use* (Seglem et al, 2015; Ystrom et al, 2014a, 2014b), *sick leave and disability pension* (Gjerde et al, 2013; Torvik et al, 2013), *resilience and self-efficacy* (Waaktaar & Torgersen, 2012, 2013) and *asthma and allergies* (Harris et al, 1997; Nystad et al, 2005).

The main aim of this paper was to outline, explicate and explain some fundamentals of twin studies. The basic twin design involves a logic carrying its own elegance, provides a scientific avenue to studying the unobserved, and enables us to estimate the total causal effects of genetic and environmental factors. In addition to elaborating on the basic principles, we have addressed some limitations of twin studies. As with other designs and scientific approaches, there are several concerns and caveats in this field. However, the major findings appear to have high validity, and a set of replicated findings has been outlined. Internationally, as well as nationally, we believe twin studies will continue to thrive – and undergo further growth alongside and as integrated with molecular genetics, psychology, medicine and epidemiology.

REFERENCES

- Ask, H., Idstad, M., Engdahl, B., & Tambs, K. (2013). Non-random mating and convergence over time for mental health, life satisfaction, and personality: The Nord-Trøndelag Health Study. *Behav Genet*, *43*(2), 108-119.
- Ask, H., Rognmo, K., Torvik, F. A., Røysamb, E., & Tambs, K. (2012). Non-random mating and convergence over time for alcohol consumption, smoking, and exercise: the Nord-Trøndelag Health Study. *Behav Genet*, *42*(3), 354-365.
- Ask, H., Waaktaar, T., Seglem, K. B., & Torgersen, S. (2016). Common etiological sources of anxiety, depression, and somatic complaints in adolescents: a multiple rater twin study. *J Abnorm Child Psychol*, *44*(1), 101-114.
- Bannink, R., Broeren, S., van de Looij-Jansen, P. M., & Raat, H. (2013). Associations between Parent-Adolescent Attachment Relationship Quality, Negative Life Events and Mental Health. *PLoS One*, *8*(11).
- Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., et al. (2011). OpenMx: An Open Source Extended Structural Equation Modeling Framework. *Psychometrika*, *76*(2), 306-317.
- Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Rev Genet*, *3*(11), 872-882.

- Canli, T., & Lesch, K. P. (2007). Long story short: the serotonin transporter in emotion regulation and social cognition. *Nature Neurosci*, *10*(9), 1103-1109.
- Carey, G. (2004). Cholesky problems. *Behav Genet*, *34*(6), 633-633.
- Czajkowski, N., Kendler, K. S., Tambs, K., Røysamb, E., & Reichborn-Kjennerud, T. (2011). The structure of genetic and environmental risk factors for phobias in women. *Psychol Med*, *41*(9), 1987-1995.
- D'Addato, A. V. (2007). Secular trends in twinning rates. *J Biosoc Sci*, *39*(1), 147-151.
- Dalgard, O. S., Bjork, S., & Tambs, K. (1995). Social Support, Negative Life Events and Mental-Health. *Br J Psychiatry*, *166*, 29-34.
- De Moor, M. H., Boomsma, D. I., Stubbe, J. H., Willemsen, G., & de Geus, E. J. (2008). Testing causality in the association between regular exercise and symptoms of anxiety and depression. *Arch Gen Psychiatry*, *65*(8), 897-905.
- De Neve, J.-E., Christakis, N. A., Fowler, J. H., & Frey, B. S. (2012). Genes, economics, and happiness. *J Neurosci Psychol Econ*, *5*(4), 193-211.
- Derks, E. M., Dolan, C. V., & Boomsma, D. I. (2006). A test of the equal environment assumption (EEA) in multivariate twin studies. *Twin Res Hum Genet*, *9*(3), 403-411.
- Donnelly, P. (2008). Progress and challenges in genome-wide association studies in humans. *Nature*, *456*(7223), 728-731.
- Flouri, E., & Kallis, C. (2011). Adverse life events and mental health in middle adolescence. *J Adolesc*, *34*(2), 371-377.
- Gjerde, L., Czajkowski, N., Røysamb, E., Orstavik, R., Tambs, K., Knudsen, G. P., et al. (2011). The heritability of cluster C personality disorders assessed by both personal interview and questionnaire. *Behav Genet*, *41*(6), 910-910.
- Gjerde, L., Røysamb, E., Czajkowski, N., Knudsen, G. P., Ostby, K., Tambs, K., et al. (2013). Personality disorders and long-term sick leave: A population-based study of young adult Norwegian twins. *Behav Genet*, *43*(6), 520-520.
- Goldman, D. (2014). The missing heritability of behavior: The search continues. *Psychophysiology*, *51*(12), 1327-1328.
- Harris, J. R., Magnus, P., Samuelsen, S. O., & Tambs, K. (1997). No evidence for effects of family environment on asthma – A retrospective study of Norwegian twins. *Am J Respir Crit Care Med*, *156*(1), 43-49.
- Harris, J. R., Magnus, P., & Tambs, K. (2006). The Norwegian Institute of Public Health twin program of research: An update. *Twin Res Hum Genet*, *9*(6), 858-864.
- Harris, J. R., Tambs, K., & Magnus, P. (1995). Sex-Specific Effects for Body-Mass Index in the New Norwegian Twin Panel. *Genet Epidemiol*, *12*(3), 251-265.
- Heath, A. C., Berg, K., Eaves, L. J., Solaas, M. H., Corey, L. A., Sundet, J., et al. (1985). Education Policy and the Heritability of Educational-Attainment. *Nature*, *314*(6013), 734-736.
- Heath, A. C., Berg, K., Eaves, L. J., Solaas, M. H., Sundet, J., Nance, W. E., et al. (1985). No Decline in Assortative Mating for Educational-Level. *Behav Genet*, *15*(4), 349-369.
- Hoyle, R. H. (2012). *Handbook of structural equation modeling*. New York: Guilford Press.
- Imaizumi, Y. (2003). A comparative study of zygotic twinning and triplet rates in eight countries, 1972-1999. *J Biosoc Sci*, *35*(2), 287-302.
- Kendler, K. S. (1993). Twin studies of psychiatric illness: Current status and future directions. *Arch Gen Psychiatry*, *50*(11), 905-915.
- Kendler, K. S., Aggen, S. H., Czajkowski, N., Røysamb, E., Tambs, K., Torgersen, S., et al. (2008). The Structure of Genetic and Environmental Risk Factors for DSM-IV Personality Disorders A Multivariate Twin Study. *Arch Gen Psychiatry*, *65*(12), 1438-1446.
- Kendler, K. S., & Baker, J. H. (2007). Genetic influences on measures of the environment: a systematic review. *Psychol Med*, *37*(5), 615-626.
- Kendler, K. S., Eaves, L. J., Loken, E. K., Pedersen, N. L., Middeldorp, C. M., Reynolds, C., et al. (2011). The impact of environmental experiences on symptoms of anxiety and depression across the life span. *Psychol Sci*, *22*(10), 1343-1352.
- Kendler, K. S., Gardner, C. O., Gatz, M., & Pedersen, N. L. (2007). The sources of co-morbidity between major depression and generalized anxiety disorder in a Swedish national twin sample. *Psychol Med*, *37*(3), 453-462.
- Kendler, K. S., Gardner, C. O., & Lichtenstein, P. (2008). A developmental twin study of symptoms of anxiety and depression: evidence for genetic innovation and attenuation. *Psychol Med*, *38*(11), 1567-1575.
- Kendler, K. S., Heath, A. C., Martin, N. G., & Eaves, L. J. (1987). Symptoms of anxiety and symptoms of depression. Same genes, different environments? *Arch Gen Psychiatry*, *44*(5), 451-457.
- Kendler, K. S., Karkowski, L. M., & Prescott, C. A. (1999). Causal relationship between stressful life events and the onset of major depression. *Am J Psychiatry*, *156*(6), 837-841.

- Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992). Major depression and generalized anxiety disorder. Same genes, (partly) different environments? *Arch Gen Psychiatry*, *49*(9), 716-722.
- Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1993). A test of the equal-environment assumption in twin studies of psychiatric illness. *Behav Genet*, *23*(1), 21-27.
- Kinderman, P., Schwannauer, M., Pontin, E., & Tai, S. (2013). Psychological Processes Mediate the Impact of Familial Risk, Social Circumstances and Life Events on Mental Health. *PLoS One*, *8*(10).
- Krishna Kumar, S., Feldman, M. W., Rehkopf, D. H., & Tuljapurkar, S. (2016). Limitations of GCTA as a solution to the missing heritability problem. *Proc Natl Acad Sci USA*, *113*(1), E61-70.
- Loehlin, J. C. (1996). The Cholesky approach: A cautionary note. *Behav Genet*, *26*(1), 65-69.
- Loehlin, J. C. (2004). *Latent variable models: an introduction to factor, path, and structural equation analysis* (4th ed. ed.). Mahwah, N.J., London: Lawrence Erlbaum Associates.
- Lykken, D., & Tellegen, A. (1996). Happiness is a stochastic phenomenon. *Psychol Sci*, *7*(3), 186-189.
- Martin, J. A., Hamilton, B. E., & Osterman, M. J. (2012). Three decades of twin births in the United States, 1980-2009. *NCHS Data Brief*(80), 1-8.
- Matteson, L. K., McGue, M. K., & Iacono, W. (2013). Is dispositional happiness contagious? The impact of the well-being of family members on individual well-being. *J Individ Differ*, *34*(2), 90-96.
- Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is Psychology Suffering From a Replication Crisis? What Does "Failure to Replicate" Really Mean? *Am Psychol*, *70*(6), 487-498.
- McGue, M. (2010). The End of Behavioral Genetics? *Behav Genet*, *40*(3), 284-296.
- McGue, M., Osler, M., & Christensen, K. (2010). Causal Inference and Observational Research: The Utility of Twins. *Perspect Psychol Sci*, *5*(5), 546-556.
- Merikangas, K. R. (1982). Assortative Mating for Psychiatric-Disorders and Psychological Traits. *Arch Gen Psychiatry*, *39*(10), 1173-1180.
- Mezuk, B., Eaton, W. W., Golden, S. H., & Ding, Y. (2008). The influence of educational attainment on depression and risk of type 2 diabetes. *Am J Public Health*, *98*(8), 1480-1485.
- Moffitt, T. E., Caspi, A., & Rutter, M. (2006). Measured Gene-Environment Interactions in Psychopathology Concepts, Research Strategies, and Implications for Research, Intervention, and Public Understanding of Genetics. *Perspect Psychol Sci*, *1*(1), 5-27.
- Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic, 1992.
- Neale, M. C., Røysamb, E., & Jacobson, K. (2006). Multivariate genetic analysis of sex limitation and G x E interaction. *Twin Res Hum Genet*, *9*(4), 481-489.
- Nes, R. B., & Røysamb, E. (2015). The heritability of subjective well-being: Review and meta-analysis. In M. Pluess (Ed.), *Genetics of Psychological Well-Being* (pp. 75-96). Oxford: Oxford University Press.
- Nes, R. B., Røysamb, E., Reichborn-Kjennerud, T., Harris, J. R., & Tambs, K. (2007). Symptoms of anxiety and depression in young adults: Genetic and environmental influences on stability and change. *Twin Res Hum Genet*, *10*(3), 450-461.
- Nes, R. B., Røysamb, E., Reichborn-Kjennerud, T., Tambs, K., & Harris, J. R. (2005). Subjective wellbeing and sleep problems: A bivariate twin study. *Twin Res Hum Genet*, *8*(5), 440-449.
- Nes, R. B., Røysamb, E., Tambs, K., Harris, J., & Reichborn-Kjennerud, T. (2006). Subjective well-being: Genetic and environmental contributions to stability and change. *Psychol Med*, *36*(7), 1033-1042.
- Nielsen, C. S., Stubhaug, A., Price, D. D., Vassend, O., Czajkowski, N., & Harris, J. R. (2008). Individual differences in pain sensitivity: Genetic and environmental contributions. *Pain*, *136*(1-2), 21-29.
- Nilsen, T. S., Knudsen, G. P., Gervin, K., Brandt, I., Røysamb, E., Tambs, K., et al. (2013). The Norwegian Twin Registry from a Public Health Perspective: A Research Update. *Twin Res Hum Genet*, *16*(1), 285-295.
- Nordsletten, A. E., Larsson, H., Crowley, J. J., Almqvist, C., Lichtenstein, P., & Mataix-Cols, D. (2016). Patterns of Nonrandom Mating Within and Across 11 Major Psychiatric Disorders. *JAMA Psychiatry*.
- Nystad, W., Røysamb, E., Magnus, P., Tambs, K., & Harris, J. R. (2005). A comparison of genetic and environmental variance structures for asthma, hay fever and eczema with symptoms of the same diseases: a study of Norwegian twins. *Int J Epidemiol*, *34*(6), 1302-1309.
- Ørstavik, R. E., Czajkowski, N., Røysamb, E., Knudsen, G. P., Tambs, K., & Reichborn-Kjennerud, T. (2014). Sex differences in genetic and environmental influences on educational attainment and income. *Twin Res Hum Genet*, *17*(6), 516-525.
- Ørstavik, R. E., Kendler, K. S., Røysamb, E., Czajkowski, N., Tambs, K., & Reichborn-Kjennerud, T. (2012). Genetic and environmental contributions to the co-occurrence of depressive personality disorder and DSM-IV personality disorder. *J Pers Disord*, *26*(3), 435-451.
- Pearson, T. A., & Manolio, T. A. (2008). How to interpret a genome-wide association study. *JAMA*, *299*(11), 1335-1344.
- Petronis, A. (2010). Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature*, *465*(7299), 721-727.

- Plomin, R. (2013). Commentary: Missing heritability, polygenic scores, and gene-environment correlation. *J Child Psychol Psychiatry*, 54(10), 1147-1149.
- Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2013). *Behavioral genetics* (Sixth edition. ed.). New York: Worth Publishers.
- Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspect Psychol Sci*, 11(1), 3-23.
- Polderman, T. J. C., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genet*, 47(7), 702-709.
- Reichborn-Kjennerud, T., Czajkowski, N., Neale, M. C., Orstavik, R. E., Torgersen, S., Tambs, K., et al. (2007). Genetic and environmental influences on dimensional representations of DSM-IV cluster C personality disorders: a population-based multivariate twin study. *Psychol Med*, 37(5), 645-653.
- Reichborn-Kjennerud, T., Czajkowski, N., Røysamb, E., Orstavik, R. E., Neale, M. C., Torgersen, S., & Kendler, K. S. (2010). Major depression and dimensional representations of DSM-IV personality disorders: a population-based twin study. *Psychol Med*, 40(9), 1475-1484.
- Reichborn-Kjennerud, T., Czajkowski, N., Ystrom, E., Orstavik, R., Aggen, S. H., Tambs, K., et al. (2015). A longitudinal twin study of borderline and antisocial personality disorder traits in early to middle adulthood. *Psychol Med*, 45(14), 3121-3131.
- Reichborn-Kjennerud, T., Stoltenberg, C., Tambs, K., Røysamb, E., Kringlen, E., Torgersen, S., & Harris, J. R. (2002). Back-neck pain and symptoms of anxiety and depression: a population-based twin study. *Psychol Med*, 32(6), 1009-1020.
- Reynolds, C. A., & Finkel, D. (2015). A Meta-analysis of Heritability of Cognitive Aging: Minding the "Missing Heritability" Gap. *Neuropsychol Rev*, 25(1), 97-112.
- Richardson, K., & Norgate, S. (2005). The equal environments assumption of classical twin studies may not hold. *Br J Educ Psychol*, 75, 339-350.
- Rietveld, C. A., Cesarini, D., Benjamin, D. J., Koellinger, P. D., De Neve, J.-E., Tiemeier, H., et al. (2013). Molecular genetics and subjective well-being. *Proc Natl Acad Sci USA*, 110(24), 9692-9697.
- Rose, R. J., & Kaprio, J. (1988). Frequency of social contact and intrapair resemblance of adult monozygotic cotwins: or Does shared experience influence personality after all? *Behav Genet*, 18(3), 309-328.
- Rose, R. J., Kaprio, J., Williams, C. J., Viken, R., & Obremski, K. (1990). Social contact and sibling similarity: Facts, issues, and red herrings. *Behav Genet*, 20(6), 763-778.
- Rutter, M. (2007). Proceeding From Observed Correlation to Causal Inference The Use of Natural Experiments. *Perspect Psychol Sci*, 2(4), 377-395.
- Rutter, M., Moffitt, T. E., & Caspi, A. (2006). Gene-environment interplay and psychopathology: multiple varieties but real effects. *J Child Psychol Psychiatry*, 47(3-4), 226-261.
- Røysamb, E., Harris, J. R., Magnus, P., Vitterso, J., & Tambs, K. (2002). Subjective well-being. Sex-specific effects of genetic and environmental factors. *Personality and Individual Differences*, 32(2), 211-223.
- Røysamb, E., Kendler, K. S., Tambs, K., Orstavik, R. E., Neale, M. C., Aggen, S. H., et al. (2011). The Joint Structure of DSM-IV Axis I and Axis II Disorders. *J Abnorm Psychol*, 120(1), 198-209.
- Røysamb, E., Nes, R. B., & Vitterso, J. (2014). Well-being: Heritable and changeable. In K. Sheldon & R. E. Lucas (Eds.), *Stability of Happiness*: Elsevier.
- Røysamb, E., Tambs, K., Reichborn-Kjennerud, T., Neale, M. C., & Harris, J. R. (2003). Happiness and health: Environmental and genetic contributions to the relationship between subjective well-being, perceived health, and somatic illness. *J Pers Soc Psychol*, 85(6), 1136-1146.
- Sadler, M., Miller, C., McGue, M., & Christensen, K. (2009). Longevity and subjective well-being: A cotwin control study. *Behav Genet*, 39(6), 678-678.
- Schooler, J. W. (2014). Metascience could rescue the 'replication crisis'. *Nature*, 515(7525), 9.
- Seglem, K. B., Waaktaar, T., Ask, H., & Torgersen, S. (2015). Genetic and Environmental Influences on Adolescents' Smoking Involvement: A Multi-informant Twin Study. *Behav Genet*, 45(2), 171-180.
- Silventoinen, K., Sarmalisto, S., Perola, M., Boomsma, D. I., Cornes, B. K., Davis, C., et al. (2003). Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Res*, 6(5), 399-408.
- Stroebe, W., & Strack, F. (2014). The Alleged Crisis and the Illusion of Exact Replication. *Perspect Psychol Sci*, 9(1), 59-71.
- Sundet, J. M., Eilertsen, D. E., Tambs, K., & Magnus, P. (1994). No Differential Heritability of Intelligence-Test Scores across Ability Levels in Norway. *Behav Genet*, 24(4), 337-339.
- Sundet, J. M., Tambs, K., Harris, J. R., Magnus, P., & Torjussen, T. M. (2005). Resolving the genetic and environmental sources of the correlation between height and intelligence: A study of nearly 2600 Norwegian male twin pairs. *Twin Res Hum Genet*, 8(4), 307-311.

- Tambs, K., Czajkowsky, N., Røysamb, E., Neale, M. C., Reichborn-Kjennerud, T., Aggen, S. H., et al. (2009). Structure of genetic and environmental risk factors for dimensional representations of DSM-IV anxiety disorders. *Br J Psychiatry*, *195*(4), 301-307.
- Tambs, K., Harris, J. R., & Magnus, P. (1995). Sex-Specific Causal Factors and Effects of Common Environment for Symptoms of Anxiety and Depression in Twins. *Behav Genet*, *25*(1), 33-44.
- Tambs, K., Kendler, K. S., Reichborn-Kjennerud, T., Aggen, S. H., Harris, J. R., Neale, M. C., et al. (2012). Genetic and environmental contributions to the relationship between education and anxiety disorders – a twin study. *Acta Psychiatr Scand*, *125*(3), 203-212.
- Tambs, K., Moum, T., Eaves, L. J., Neale, M. C., Midthjell, K., Lund-Larsen, P. G., & Naess, S. (1992). Genetic and environmental contributions to the variance of body height in a sample of first and second degree relatives. *Am J Phys Anthropol*, *88*(3), 285-294.
- Tambs, K., Ronning, T., Prescott, C. A., Kendler, K. S., Reichborn-Kjennerud, T., Torgersen, S., & Harris, J. R. (2009). The Norwegian Institute of Public Health Twin Study of Mental Health: Examining Recruitment and Attrition Bias. *Twin Res Hum Genet*, *12*(2), 158-168.
- Tandberg, A., Bjorge, T., Bordahl, P. E., & Skjaerven, R. (2007). Increasing twinning rates in Norway, 1967-2004: the influence of maternal age and assisted reproductive technology (ART). *Acta Obstet Gynecol Scand*, *86*(7), 833-839.
- Tellegen, A., Lykken, D. T., Bouchard, T. J., Wilcox, K. J., Rich, S., & Segal, N. L. (1988). Personality Similarity in Twins Reared Apart and Together. *J Pers Soc Psychol*, *54*(6), 1031-1039.
- Torgersen, S., Czajkowski, N., Jacobson, K., Reichborn-Kjennerud, T., Røysamb, E., Neale, M. C., & Kendler, K. S. (2008). Dimensional representations of DSM-IV cluster B personality disorders in a population-based sample of Norwegian twins: a multivariate study. *Psychol Med*, *38*(11), 1617-1625.
- Torvik, F. A., Gjerde, L., Knudsen, G. P., Reichborn-Kjennerud, T., Røysamb, E., Czajkowski, N., et al. (2013). Anxiety, depression and sick leave due to various diagnoses: A population-based study of young adult Norwegian twins. *Behav Genet*, *43*(6), 544-544.
- Tsankova, N., Renthal, W., Kumar, A., & Nestler, E. J. (2007). Epigenetic regulation in psychiatric disorders. *Nature Rev Neurosci*, *8*(5), 355-367.
- Turkheimer, E. (2012). Genome-wide complex trait analysis and the missing heritability problem. *Behav Genet*, *42*(6), 972-972.
- Vassend, O., Røysamb, E., & Nielsen, C. S. (2011). Dental anxiety in relation to neuroticism and pain sensitivity. A twin study. *J Anxiety Disord*, *25*(2), 302-308.
- Vassend, O., Røysamb, E., & Nielsen, C. S. (2013). Five-factor personality traits and pain sensitivity: A twin study. *Pain*, *154*(5), 722-728.
- Waaktaar, T., & Torgersen, S. (2012). Genetic and environmental causes of variation in trait resilience in young people. *Behav Genet*, *42*(3), 366-377.
- Waaktaar, T., & Torgersen, S. (2013). Self-Efficacy Is Mainly Genetic, Not Learned: A Multiple-Rater Twin Study on the Causal Structure of General Self-Efficacy in Young People. *Twin Res Hum Genet*, *16*(3), 651-660.
- Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., et al. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet*, *46*(11), 1173-1186.
- Wright, S. (1921). Correlation and causation. *J Agric Res*, *20*, 557-585.
- Wright, S. (1934). The method of path coefficients. *Ann Math Stat*, *5*(3), 161-215.
- Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat Genet*, *42*(7), 565-569.
- Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2013). Genome-wide complex trait analysis (GCTA): methods, data analyses, and interpretations. *Methods Mol Biol*, *1019*, 215-236.
- Ystrom, E., Kendler, K. S., & Reichborn-Kjennerud, T. (2014a). Early age of alcohol initiation is not the cause of alcohol use disorders in adulthood, but is a major indicator of genetic risk. A population-based twin study. *Addiction*, *109*(11), 1824-1832.
- Ystrom, E., Reichborn-Kjennerud, T., Neale, M. C., & Kendler, K. S. (2014b). Genetic and Environmental Risk Factors for Illicit Substance Use and Use Disorders: Joint Analysis of Self and Co-twin Ratings. *Behav Genet*, *44*(1), 1-13.