TwinLife Working Paper Series

No. 01, April 2017

Zygosity Determination in Twin Studies

A Validation of Zygosity Questionnaires Using DNA in the German TwinLife Study

by Franziska Lenau, Elisabeth Hahn, Anna-Lena Peters, Juliana Gottschling, Wolfgang Thiel, & Frank M. Spinath

www.twin-life.de/en/twinlife-working-paper-series

F. Lenau, E. Hahn, A.-L. Peters, J. Gottschling, W. Thiel, & F. M. Spinath

Zygosity Determination in Twin Studies

TwinLife Working Paper Series No. 01
Project TwinLife “Genetic and social causes of life chances”
Bielefeld, April 2017

TwinLife Working Paper Series No. 01
General Editors: Martin Diewald, Rainer Riemann and Frank M. Spinath
ISSN 2512-4048

This publication has been funded by the German Research Foundation (DFG).

TwinLife Working Papers are refereed scholarly papers. Submissions are reviewed by the general editors before a final decision on publication is made.

The Working Paper Series is a forum for presenting works in progress. Readers should communicate comments on the manuscript directly to the author(s).

The papers can be downloaded from the project website:
http://www.twin-life.de/en/twinlife-working-paper-series

TwinLife “Genetic and social causes of life chances”
University of Bielefeld
Faculty of Sociology
PO Box 100131
D-33501 Bielefeld
Germany

Phone: +49 (0)521 106-4309
Email: martin.diewald@uni-bielefeld.de
Web: http://www.twin-life.de/en

Universität Bielefeld

UNIVERSITÄT DES SAARLANDES

Zygosity Determination in Twin Studies: A Validation of Zygosity Questionnaires Using DNA in the German TwinLife Study

Franziska Lenau[1], Elisabeth Hahn[1], Anna-Lena Peters[2], Juliana Gottschling[1], Wolfgang Thiel[2], and Frank M. Spinath[1]

[1] Department of Psychology, Saarland University, Saarbruecken, Germany
[2] Department of Psychology, Bielefeld University, Bielefeld, Germany

Zygosity determination using similarity ratings is frequently applied in twin studies. A correct determination of zygosity is essential for the estimation of heritability and environmental influences on phenotypes. Therefore, the present study examined the validity of two similarity questionnaires used in the German *TwinLife* study, in which data from 4,097 twin pairs and their families were assessed: twin children's zygosities were determined with the *Zygosity Questionnaire for Young Twins*, which was administered in parent-report form. For adolescent twins, the *Self Report Zygosity Questionnaire* was used. For the present validation analyses, DNA samples of *N* = 328 twin pairs were collected via buccal swabs. In this DNA subsample, questionnaires were filled out by parents for *n* = 212 (aged 4 to 12 years) twin pairs while self-reports were collected from *n* = 116 adolescent twins (16 to 23 years of age). Using DNA-based zygosity as criteria, correct classification rates of 97% for parent- and 92% for self-reports were established and cross-validated. Additionally, classification rates based on a single item and variants of questionnaire based zygosity determination used in other twin studies were calculated and compared. Implications of incorrectly classified zygosity on genetic and environmental estimates in twin studies are discussed.

Keywords: Zygosity Determination, Behavior Genetics, Similarity Questionnaire, Twin Studies

Correspondence concerning this article should be addressed to Franziska Lenau, Department of Psychology, Saarland University, Germany, E-mail: franziska.lenau@uni-saarland.de

The determination of zygosity (i.e., monozygotic vs. dizygotic) is a key element in behavior genetic twin studies. The rationale for estimating the relative influences of genetic and environmental factors on phenotypes is usually based on the comparison of monozygotic (MZ) and dizygotic (DZ) twin similarity and their corresponding genetic similarity. Correct classifications of twin pairs as MZ or DZ are essential, because misclassifications can lead to a systematic bias in parameter estimates. If, for example, MZ twin pairs are misclassified as DZ twins, genetic influences may be underestimated whereas common environmental effects may be overestimated (Conley et al., 2013). In general, two main approaches for zygosity determination can be distinguished: one that uses *biological* information and one that uses *physical* characteristics.

Biological characteristics such as blood type, serological markers (antibodies in the blood serum) or genetic markers (easily identifiable short DNA sequences) can be used to determine the zygosity of twin pairs (Song et al., 2010). This method draws on the differences in genetic similarities of MZ and DZ twins: MZ twins are mostly[1] genetically identical, whereas DZ twins share, on average, 50% common genes (e.g., Boomsma et al., 1999; McGue et al., 2010). Biological methods use this difference to establish zygosity by inspecting the co-twins of a pair on a number of highly polymorphic alleles. If no differences are found for a predefined number of loci, twins are classified as MZ. Also, thresholds can be used, such as a minimum number of differences (e.g., two different alleles out of five) to classify twins as DZ (see Becker et al., 1997) with thresholds typically depending on the number of alleles examined. Utilization of such highly polymorphic genetic markers usually yield correct classification rates close to 100% (Becker et al., 1997). Therefore, this method is the most reliable and valid variant of zygosity determination. To collect and analyze DNA of each twin participant in order to reach the best classification rate possible may be practicable in small samples; but the larger the sample, which is desirable in twin studies to increase statistical power, the less practical and the more cost-intensive is the collection of DNA data. Moreover, the necessity to collect blood or buccal swabs for the extraction of DNA can reduce the willingness to participate in a study (Spitz et al., 1996). Participants may be deterred by the invasive procedure, or may have concerns with respect to data privacy. Furthermore, for some methods (such as blood samples) participants and researchers have to meet in person, which can be difficult to realize in studies that are designed as pure questionnaire or online studies. Finally, DNA analyses are still cost-intensive; even though the price per one pair of DNA samples has decreased considerably during the last three decades, as of yet it still can cost more than $100 per pair of DNA samples.

A second method that has frequently been used in twin studies and that has advantages with regard to financial considerations and practicability involves the analysis of *physical* characteristics such as fingerprints, similarities of face and body features as well as information from the twins' developmental courses (Becker et al., 1997; Song et al., 2010). The underlying logic of this method is, that because of their genetic match, MZ twins show on average greater physical similarities than DZ twins (Martin & Martin, 1975). Therefore, similarity ratings can be leveraged to determine zygosity. Because twins with a greater physical similarity are probably confused by other persons (e.g., parents, teachers, or friends) more often, the frequency of twin confusions can also be utilized. In contrast to biological characteristics, physical characteristics can be assessed not only through self-reports (e.g., for adolescent and adult twins) but also via parent- or peer-reports (e.g., teachers) on young twins.

Similarity questionnaires have some advantages in comparison to the collection of DNA: they are non-invasive, feasible as paper-pencil as well as online questionnaires, and in consequence, inexpensive. However, correct classification rates are lower for similarity questionnaires compared to DNA based zygosity

[1] Even though MZ twins are often called genetically identical, they could differ in DNA, caused by DNA mutations or epi-genetic modifications (e.g., Gringras & Chen, 2001). These differences are likely to be very small, if at all existent.

determination, even though the majority of studies report correct classification rates above 90% for different questionnaires (e.g., Chen et al., 1999; Spitz et al., 1996). The lower correct classification rates – compared to DNA analyses – can, in part, be explained by the fluent transition of similarity between DZ twins based on the proportion of common genes, which is only on average 50% (van Dongen et al., 2012; Visscher et al., 2007). Therefore, some DZ twin pairs may have greater genetic similarities than 50% whereas others are more dissimilar. Especially DZ twins with a greater physical resemblance may be at a greater risk of mis-classification.

As a correct classification of zygosity is essential for a valid estimation of heritability, zygosity questionnaires need to be validated through DNA-based zygosity determination: correct classification rates can be calculated as the percentage of twins that are classified into the same zygosity group when using questionnaire data as by using DNA data. Due to sample specificity, "classification is usually best when applied to the sample on which the classification formulae or rules were developed" (Jackson et al., 2001, p. 12). Therefore, it is advisable to provide additional evidence for the validity of the questionnaire based on other samples (e.g., via cross-validation), to provide evidence for its generalizability. As the physical similarity of twins, e.g., regarding height or weight, within a pair can change over time – MZ twins usually remain highly similar while DZ twins tend to become less similar over time (Åkerman & Fischbein, 1992) – it is also important to validate questionnaires in different age groups.

Most large scale twin studies use questionnaires for zygosity determination, however, the number of questions, and in consequence the number of physical characteristics addressed, differs substantially. For example, two self-reported questions – one addressing the overall physical similarity of twins (called '*peas in a pod*'-question) and the other addressing the frequency of confusion by people meeting them for the first time – were used in the *Swedish Twin Registry* (Lichtenstein et al., 2002) for adult twins. Three characteristics, namely the '*peas in a pod*'- and the '*mistaken by people meeting first time*'-question as well as parents' belief about zygosity were used in the *Virginia Twin Study* (Eaves et al., 1997). Other studies, such as the *Michigan State University Twin Registry* (Burt & Klump, 2013) and the *Vietnam Era Twin Registry* (Xian et al., 2000) use zygosity questionnaires with up to 19 characteristics. Typically, longer and hence more reliable scales are associated with higher validity, which is why we include a comparison of zygosity scales of different length in this study.

The present research is realized as part of the German *TwinLife Study* (Diewald et al., 2016) on the development of social inequalities (for an overview see Hahn et al., 2016). Zygosity was assessed via a parent-report questionnaire in young twins (Goldsmith, 1991) and via a self-report questionnaire for adolescents (Onisczenko et al., 1993). DNA samples were collected in a subsample of child and adolescent twins to enable the validation of the zygosity questionnaires. To this end, the accuracy of the classification based on the results of questionnaire data was evaluated in comparison to DNA data by using discriminant functions. Next, cross-validation analyses were carried out for both, self- and parent-report questionnaires. Also, correct classification rates based on single characteristics (as realized in other twin studies) were investigated compared to the full questionnaire. In addition, heritability estimates resulting from different zygosity determination methods – and therefore possibly based on different groups of twins classified as MZ and DZ – were compared to illustrate the importance of solid zygosity determination.

Method

Participants

The present report used data from the German *TwinLife Study* (Hahn et al., 2016), a genetically informative, longitudinal study addressing the development of social inequality. In *TwinLife*, a total of 4,097 MZ and DZ same-sex twin pairs, encompassing four different age cohorts and their families, were assessed in the first wave of the study and will be followed over a time span of 9 years. Data collection started in September 2014 with the first face-to-face interview in the participants' households. For the first measurement occasion, families were invited to participate when the twins were 5, 11, 17, and 23 years old.[2] The present study reports results from the first half of the sample (i.e., 2,009 twins and their families collected between September 2014 and May 2015), in which the DNA sampling was realized. As described above, similarity questionnaires were addressed to all twin pairs. In addition to the questionnaire data, we collected DNA samples using buccal swabs in a subsample of N = 328 twin pairs: n = 107 (cohort 1; 54% female; M_{age} = 5.0), n = 105 (cohort 2; 57% female; M_{age} = 11.0), n = 116 (cohort 3; 58% female; M_{age} = 17.3). For the younger twin groups (cohorts 1 and 2), similarity reports provided by one parent were used in the present analyses. If parent-reports were completed by both parents (1%), only the data provided by mothers were analyzed to harmonize the source of the information. Altogether, we used mother reports in 82%. In the remaining 18% of the cases, mother reports were not available; therefore, we used father reports. For the older twin group, both twins filled in the questionnaire themselves.

[2] The *TwinLife* sample consists of four age cohorts (5, 11, 17, and 23 years at the first measurement occasion). Each cohort encompasses birth cohorts spanning 2 years due to the small number of expected twin births in Germany. To assess each twin family within the same cohort at about the same age, each wave is organized into two half-waves, each half following the respective birth cohort of twins. For example, families with 5-year-old twins born in 2009 were studied in 2014 and families with 5-year-old twins born in 2010 were studied in 2015.

Zygosity Questionnaires

The *Zygosity Questionnaire for Young Twins* (Goldsmith, 1991) developed for parent-report, was used for the determination of zygosity in 1,011 young twins[3] (cohorts 1 and 2; aged 4 to 12 years). The *Self Report Zygosity Questionnaire* by Onisczenko and colleagues (1993) was used for zygosity determination in 991 adolescent twin pairs[4] (cohorts 3 and 4; aged 16 to 23 years). Some characteristics, e.g., height, were assessed separately for both twins. Other characteristics were assessed with a single item for the pair (e.g., differences in hair color). The content of both, parent- and self-report questionnaire can be categorized into three areas: the first part contained questions regarding *physical similarities*. All participants were asked about the twins' height, differences in hair texture, differences in eye color, and differences in ear lobes, blood types, and rhesus factors. Parents were additionally asked about the twins' differences in hair color, differences regarding first teeth, similarity as the twins grew older, and 'overall' physical similarity.[5] In the self-report questionnaire the older twins were additionally asked about their overall hairiness, eye colors, tendency to sweat, skin color, and similarity in the frequency of sickness in childhood.

The second part of both questionnaires contained items regarding *confusion*, in which parents were asked if they could correctly identify each twin in a photograph at the age of two to 4 years, whether twins were ever mistaken when together, and whether they were mistaken by people meeting them for the first time, by babysitter or daycare workers, by the other parent, by older siblings, by close and casual friends as well as other relatives. In the self-report version, twins were asked whether others could tell them apart on recent photographs, and whether they were ever mistaken by another. If this last question was

[3] The zygosity questionnaire was not completed for 5 out of 1,016 twin pairs.

[4] The zygosity questionnaire was not completed for 2 out of 993 twin pairs.

[5] In the literature this question is called the '*peas in the pod*'-question, as one of the answer options is '*twins are as alike as two peas in a pod*'.

affirmed, twins were additionally asked whether they were mistaken by people meeting them for the first time, by teachers, by parents, by siblings and by friends. Furthermore, twins were asked whether parents had to make particular effort to tell them apart.

In the third part of the questionnaires parents and twins were asked whether they considered themselves mono- or dizygotic and if they were ever told by medical staff about the twins' zygosity.

In the young as well as the adult twin sample, difference scores between twins of a pair were calculated for each characteristic of the questionnaire.[6] On average, within-pair difference scores on the zygosity questionnaire items should be smaller for MZ than for DZ twin pairs.

DNA Genotyping

DNA samples from each twin were collected by the *TwinLife* interviewers using buccal swaps. Accordingly, interviewers were trained and equipped with a declaration of consent as well as test tubes. If preferred, parents or the older twins themselves could carry out the DNA collection under the guidance of the interviewer. DNA was extracted with the chelating agent Chelex®, which is based on a cook-lysis technique. Afterwards, it was amplified per *Polymerase Chain Reaction* (PCR) with the STR-Kit PowerPlex21 by Promega. With *capillary electrophoresis* (CE), 21 Short Tandem Repeat markers (STR) were visualized in the electropherogram. These STR's included all 13 Combined DNA Index System (CODIS) loci (D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, CSF1PO, FGA, TH01, TPOX, vWA), plus eight additional loci (Amelogenin, Penta D, Penta E, D1S1656, D2S1338, D6S1043, D12S391 and D19S433), yielding a total of 42 alleles. Twin pairs with differences in none up to two alleles were classified as monozygotic. If four or more alleles differed, twins were classified as dizygotic. Exactly three differing alleles would have resulted in additional analyses; however, this did not occur in the present analyses. A similar procedure was described in detail by Becker et al. (1997).

Analyses

Accuracy of Zygosity Determination by Discriminant Function. Zygosity determined by DNA genotyping was used as the true criterion. By using the true classification of each twin pair, two discriminant functions were calculated including beta weights for all items assessed in the questionnaires: one function for the younger groups (parent-report) and one for the older group (self-report). Based on these discriminant functions, zygosity was calculated for the two groups again. *Correct classification rates* were established as the percentage of twins that were classified in accordance with DNA results using questionnaire data in the discriminant function.

Cross-Validation. Different approaches were used for cross-validation of the zygosity determination functions in the parent- and the self-report sample: To provide validation for the parent-report questionnaire and the corresponding discriminant function developed in the present data, the sample was randomly split into two cross-validation samples ($n_{1.Half}$ = 105; $n_{2.Half}$ = 107). In each of these samples, a separate determination function was established and was then used in the other half-sample to compare the resulting classification rates. For the self-report questionnaire, the discriminant function developed in the present data was compared to a function developed and already validated by DNA in another sample of twins derived from the *Bielefeld Longitudinal Study of Adult Twins* (*BiLSAT*; Kandler et al., 2013). The '*TwinLife*' function was applied to the *BiLSAT* data and vice versa to cross-validate the self-report questionnaire and the respective determination functions.

Prediction Accuracy. To provide further validation, *discriminant function analysis* was used to determine which items discriminated best between MZ and DZ twins. For this purpose, *prediction accuracy* per item was calculated as the

[6] For a detailed documentation of zygosity data in the *TwinLife* study (incl. SPSS scripts) see Lenau & Hahn (2017).

percentage of twin pairs whose zygosity was correctly identified based on the responses to the respective single item. To compare our results with common practice strategies, we also calculated zygosity in accordance with other twin studies that used two ('*peas in a pod*'- and the '*mistaken by people meeting first time*') or three items ('*peas in a pod*'- and the '*mistaken by people meeting first time*'-question as well as parents belief about zygosity). As the '*peas in a pod*'-question is part of both variants, but is included in the present study only in the parent-report questionnaire, the according analyses were carried out only for the parent-reports.

Intraclass Correlations (ICC's). To demonstrate consequences of differing accuracy of zygosity determination, twin similarity based on ICC's for *cognitive ability* were calculated for different variants of zygosity determination, namely: DNA (1), full discriminant function (2), two-item (3) and three-item determination method (4) as well as parents / own belief (5). Cognitive ability was measured via three (children aged younger than 10 years), respectively four (all persons aged 10 years and older) subtests of the *Culture Fair Test* (CFT; Weiß, 2006; Weiß & Osterland, 2012), a widely used and well validated cognitive test battery, that captures non-verbal (fluid) intelligence as a proxy for general cognitive ability.[7]

Results

DNA Diagnosis of Twin Zygosity

DNA determination for the 107 same-sex twin pairs aged 4 to 6 years (M_{age} = 5.0 years), revealed that 34 were DZ and 73 were MZ. In the subsample of the 105 same-sex twin pairs aged 10 to 12 years (M_{age} = 11.0 years), 35 were DZ and 70 were MZ. Among the 116 same-sex adolescent twins (M_{age} = 17.3 years), 36 were DZ and 80 were MZ.

Accuracy of Zygosity Determination by Discriminant Function

The discriminant function developed in the present parent-report sample resulted in an overall correct classification rate of 97%, with identical accuracy for MZ and DZ twins. Overall, correct classification rates were a little higher for males (99%) than for females (95%), and 5-year-old (98%) compared to 11-year-old twins (95%).

The discriminant function developed in the present self-report sample, resulted in an overall correct classification rate of 96%. MZ twins were classified with a higher accuracy (98%) than DZ twins (92%). There was no difference in the correct classification of male (96%) and female (96%) twins.

Cross-Validation

To validate the discriminant function[8] for the parent-report questionnaire, the sample was randomly split into two halves ($n_{1.Half}$ = 105; $n_{2.Half}$ = 107) and a separate discriminant function was developed in each of these samples. Classification rates by using functions 'cross-wise' in the respectively other half ranged from 81% to 97%, as shown in TABLE 1. The one relatively low correct classification rate occurred for DZ twins when the function, developed in the first half was used in the second half. All remaining correct classification rates were above 90%.

TABLE 1: *Cross Validation*

			Correct classification in %		
	Function	***Sample***	***MZ***	***DZ***	***Overall***
Parent-Report	*1. Half* [1]	*1. Half* [1]	98.6	96.9	98.1
	2. Half [2]	*2. Half* [2]	92.9	97.3	94.4
	2. Half [2]	*1. Half* [1]	95.9	96.9	96.2
	1. Half [1]	*2. Half* [2]	95.7	81.1	90.7
Self-Report	*TwinLife* [o]	*TwinLife* [o]	97.5	91.7	95.7
	BiLSAT [+]	*BiLSAT* [+]	97.2	88.5	93.5
	BiLSAT [+]	*TwinLife* [o]	90.0	97.2	92.2
	TwinLife [o]	*BiLSAT* [+]	97.9	83.7	91.8

NOTES:
[1] First half of random sample split (n = 105); [2] Second half of random sample split (n = 107); [+] *BiLSAT* sample with N = 245 twin pairs; [o] TwinLife sample with n = 212 twin pairs

[7] For a more detailed description of the assessment of and handling with cognitive ability in the *TwinLife* study see Gottschling (2017).

[8] This function uses 15 of 21 characteristics. Due to little predictive power, the following characteristics were not used for zygosity determination: difference height, difference blood type, difference Rhesus factor, mistaken on actual photograph, mistaken by close friends, information by medical staff.

TABLE 2: *Accuracy of Zygosity Determination by Single Characteristics*

	Accuracy in % of Parent-Report				Accuracy in % of Self-Report			
Question	**MZ**	**DZ**	**All**	***n*** [#]	**MZ**	**DZ**	**All**	***n*** [#]
Difference height[x]	80.4	46.9	69.8	202	85.1	60.0	77.9	104
Difference hair texture	69.9	88.4	75.9	212	56.3	91.7	67.2	116
Difference eye color	95.8	63.8	**85.4**	212	96.3	57.6	85.0	113
Difference ear lobes	84.6	81.2	83.5	212	68.8	75.0	70.7	116
Difference blood type[x]	**100.0**	40.0	83.3	18	**100.0**	28.6	75.0	20
Difference Rhesus factor[x,y]	**100.0**	**7.7**	69.2	39	**100.0**	**0.0**	62.5	16
Difference hair color [a] / Difference hairiness[b]	83.9	85.5	84.4	212	95.0	19.4	71.6	116
Teeth at the same time[a] / Difference eye color[b,z]	70.6	58.0	66.5	212	75.0	83.3	77.6	116
Similarity as grown older[a] / Difference sweating[b]	55.9	62.3	**58.0**	212	90.9	35.3	73.9	111
Peas in a pod[a] / Difference skin color[b]	81.1	94.2	**85.4**	212	98.8	2.8	69.0	116
Difference sickness[b]					58.7	63.6	60.2	108
Mistaken on photograph[x]	77.6	85.5	80.2	212	37.5	97.2	56.0	116
Mistaken when together[a] / Ever mistaken[b]	76.9	89.9	81.1	212	97.5	65.7	87.8	115
Mistaken by people meeting first time	93.0	51.6	80.4	204	88.8	88.6	**88.7**	115
Mistaken by babysitter[a] / teacher[b]	88.7	67.7	81.7	93	93.8	68.6	86.1	115
Mistaken by parent	54.4	92.0	64.5	186	26.3	91.4	46.1	115
Mistaken by sibling	**49.3**	**100.0**	63.7	102	**10.0**	**100.0**	**37.4**	115
Mistaken by (close) friends[x]	69.3	88.5	75.1	201	57.5	85.7	66.1	115
Mistaken by casual friends[a] / Effort to keep apart[b]	92.1	62.9	83.3	201	26.9	97.2	49.1	114
Mistaken by other relatives[a]	75.7	80.0	77.0	200				
Own belief[y]	75.2	98.5	83.3	186	81.4	**100.0**	88.3	94
Information medical staff[x,y]	64.5	**100.0**	78.0	100	70.8	**100.0**	83.3	42

Notes:
[#] Not each question was answered by all parents or twins, therefore *n* fluctuates; [a] Question only asked in parent-report questionnaire; [b] Question only asked in self-report questionnaire; [x] Question not used for zygosity determination in 5-year-old twins; [y] Question not used for zygosity determination in 17-year-old twins; [z] calculated from eye colors

To cross-validate the determination function [9] for the self-report questionnaire, a function [10], which was developed and already validated by DNA in *BiLSAT* (Kandler et al., 2013) was used. Using the newly developed '*TwinLife*' function and the '*BiLSAT*' function 'cross-wise' in the other sample, respectively, resulted in correct classification rates ranging from 84% to 98% (see TABLE 1). The *existing* function worked well in the present sample (altogether 92% correct classifications). As this function was developed in the *BiLSAT* sample, this was an unbiased result with regard to specific sample characteristics. Therefore, this determination function shows satisfying generalizability.

Based on these accuracies and cross-validation results in the subsamples of twin pairs, zygosity was determined in the whole *TwinLife* study via the *newly developed* function in the sample of young twins (n = 2,048) and via the '*BiLSAT*' function in the sample of adolescent twins (n = 2,042).

Accuracy of Zygosity Determination by Single Attributes

The classification rates based on single attributes are displayed in TABLE 2. They ranged from 58% ('similarity *as grown older*') to 85% ('*difference in eye color*' or '*peas in a pod*') in the parent-report, and between 37% ('*mistaken by sibling*') and 89% ('*mistaken by people meeting first time*') in the self-report sample.

Comparison of common practice strategies. In the parent-report sample, using two questions ('*peas in a pod*' and '*mistaken by people meeting first*

[9] The newly developed function uses 10 of 21 characteristics. As they showed non-significant predictive values, the following characteristics were not used for zygosity determination in this function: difference blood type, difference Rhesus factor, difference skin color, difference sickness, mistaken on new photograph, ever mistaken, mistaken by parents, mistaken by sibling, mistaken by friends, own belief, information medical staff.

[10] This function, developed in *BiLSAT*, uses 17 of the assessed 21 characteristics for zygosity determination. Not used are Rhesus factor, mistaken by friends, own belief and information by medical staff.

time') for zygosity determination yielded an overall correct classification rate of 85%. Using one additional question ('own *belief*') resulted in an overall correct classification rate of 92%.

Intraclass Correlations for Cognitive Ability

Intraclass correlations were calculated for cognitive ability adjusted for age and sex. In the youngest sample of 5-year-old twins ICC's by using DNA for zygosity determination was .61 for MZ and .55 for DZ twins. The biggest discrepancy from these values was observed when parents' own beliefs about zygosity were used: ICC's changed to .67 for MZ (difference = +.06) and .52 for DZ (difference = –.03) twin pairs. In the sample of 11-year-old children, ICC's of .52 for MZ and .29 for DZ twins were found by using DNA determined zygosity. The largest discrepancy from these values was observed for the zygosity determination using three items: ICC's changed to .40 for MZ (differ ence = –.12) and .40 for DZ (difference = +.11) twins. In the adolescent sample ICC's were .65 for MZ and .34 for DZ twin pairs by using DNA for zygosity determination. For MZs, hardly any discrepancies were observed (ICC = .63, difference = –.02) when using the discriminant function, for DZ the largest difference was found when twins' beliefs were used for zygosity determination (ICC = .41, difference = +.07).

In the youngest cohort, the use of inferior zygosity information lead to an underestimation of shared environmental and an overestimation of genetic influences. In the 11-year and adolescent cohorts, the pattern changed in a way that using inferior zygosity information primarily lead to a pronounced underestimation of genetic influences.

Discussion

Even though DNA genotyping produces nearly perfect classification rates of twins' zygosity, it comes along with a higher burden for participants (e.g., collecting DNA from buccal swaps or blood) as well as substantial costs, especially for large twin studies. In the German *TwinLife* study, we employed similarity questionnaires to determine zygosity and found high correct classification rates compared to DNA results between 95% and 99% for parent-report, and between 90% and 97% in self-ratings.

Cross-validation was used to evaluate the generalizability of the zygosity determination functions. High correct classification rates suggested that it is feasible to use the determination function, which was developed in the present subsample, for zygosity determination in the whole parent-report sample of the *TwinLife* study (n = 2,048). In the self-report sample of adolescent twins, cross-validation was implemented using data from *BiLSAT* (Kandler et al., 2013). High correct classification rates were found for the established zygosity determination function developed in *BiLSAT* and applied to the *TwinLife* sample of adult twin pairs (n = 2,042).

Accuracies for single characteristics ranged from 58% to 85% in the parent-report sample and between 37% and 89% in the self-report sample. In sum, some single characteristics revealed comparatively high classification rates in the prediction of zygosity. However, most items were adequate only for the correct classification of either MZ or DZ twins. Furthermore, two additional variants, of zygosity determination were tested in the parent-report sample. For these variants, classification rates were satisfying: accuracy rose from 85% (two items) to 92% (three items). Zygosity determination based on a combination of several characteristics generally yielded increased overall correct classification rates, and also contributed to fewer systematic differences in the classification of MZ and DZ twins. This is important, because misclassifications of MZ and DZ twins have different consequences: as already mentioned misclassification of MZ twins as DZ results in an underestimation of genetic, and an overestimation of common environmental effects. A misclassification of DZ twins as MZ has the opposite effect.

Intraclass correlations (ICC's) for cognitive ability were calculated separately for different variants of zygosity determination to demonstrate possible effects of incorrect classifications. Absolute differences ranged from .00 to .12 in

different directions when ICC's resulting from zygosity determination via DNA were compared with other determination methods. One particularly striking change in the resulting genetic and environmental estimates occurred when the three-item variant was used in 11-year-old twins, since the inferior zygosity information caused the heritability estimate to drop from .46 to zero, that is the complete *absence* of genetic effects.

Limitations and future directions. Some limitations of the present study need to be addressed in the following. First, there were about twice as many MZ than DZ twins in each of the three age groups. This may well have been the result of direct or indirect selective participation in the DNA collection process since participants were invited by the interviewers to the DNA collection. It is possible that interviewers asked MZ more often than DZ twins. Possibly, they did not see the need to ask very dissimilar DZ twin pairs. Another possible explanation is that MZ twins gave their consent more often than DZ twins, for example because they may be more interested or more often uncertain about their zygosity compared to DZ twins. Some credit to the latter explanation appears to come from a second look at the '*own belief*'-question because correct classification rates were markedly lower for MZ than for DZ twins. A further limitation of this study is that it was not possible for the parent-report sample of 5- and 11-year-old twins to perform a 'true' cross-validation. Instead, we had to rely on the random sample split method.

Conclusion. If it is not possible to collect DNA for each twin pair in a twin study and it is therefore necessary to use questionnaire-based zygosity determination methods, it is advisable to assess several physical characteristics of the twins. Utilizing zygosity questionnaires of this kind has two major benefits: first, overall high correct classification rates can be achieved. Second, correct classification rates of MZ and DZ twins do not indicate systematic bias when several characteristics are used.

Conflict of Interest and Funding Statements

This research was supported by a grant from the German Research Foundation awarded to Martin Diewald (DI 759/11-1), Rainer Riemann (RI 595/8-1), and Frank M. Spinath (SP 610/6-1). DNA genotyping was conducted by the *Eurofins Medigenomix Forensik GmbH*. The authors declared no potential conflicts of interest with respect to the research, authorship, and publication of this article.

References

Åkerman, B. A., & Fischbein, S. (1992). Within-pair Similarity in MZ and DZ Twins from Birth to Eighteen Years of Age. *AMG Acta Geneticae Medicae et Gemellologiae: Twin Research*, *41*, 155–164. doi: 10.1017/S0001566000002361

Becker, A., Busjahn, A., Faulhaber, H., Bähring, S., Robertson, J., Schuster, H., & Luft, F. (1997). Twin zygosity. Automated determination with microsatellites. Journal of Reproductive Medicine, 42, 260–266.

Boomsma, D. I., de Geus, E. J. C., van Baal, G. C. M., & Koopmans, J. R. (1999). A religious upbringing reduces the influence of genetic factors on disinhibition: Evidence for interaction between genotype and environment on personality. Twin Research, 2, 115–125. doi: 10.1375/136905299320565988

Burt, S. A., & Klump, K. L. (2013). The Michigan State University Twin Registry (MSUTR): an update. *Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies*, *16*, 344–350. doi: 10.1017/thg.2012.87

Chen, W. J., Chang, H. W., Wu, M. Z., Lin, C. C., Chang, C., Chiu, Y. N., & Soong, W. T. (1999). Diagnosis of Zygosity by Questionnaire and Polymarker Polymerase Chain Reaction in Young Twins. Behavior Genetics, 29, 115–123. doi: 10.1023/A:1021660506222

Conley, D., Rauscher, E., Dawes, C., Magnusson, P. K. E., & Siegal, M. L. (2013). Heritability and the Equal Environments Assumption: Evidence from Multiple Samples of

Misclassified Twins. Behavior Genetics, 43, 415–426. doi: 10.1007/s10519-013-9602-1

Diewald, M., Riemann, R., Spinath, F. M., Gottschling, J., Hahn, E., Kornadt, A. E., Schulz, W., Schunck, R., Baier, T., Bartling, A., Kaempfert, M., Krell, K., Lang, V., Lenau, F., Nikstat, A., & Peters, A.-L. (2016). TwinLife. *GESIS Data Archive, Cologne, ZA6701 Data file Version 1.0.0*. doi: 10.4232/1.12665

Eaves, L. J., Silberg, J. L., Meyer, J. M., Maes, H. H., Simonoff, E., Pickles, A., Rutter, M., Reynolds, C. A., Heath A. C., Truett, K. R., Neale, M. C., Erikson, M. T., Loeber, R., & Hewitt, J. K. (1997). Genetics and Developmental Psycho-pathology: 2. The Main Effects of Genes and Environment on Behavioral Problems in the Virginia Twin Study of Adolescent Behavioral Development. Journal of Child Psychology and Psychiatry, 38, 965–980. doi: 10.1111/j.1469-7610.1997.tb01614.x

Goldsmith, H. H. (1991). A Zygosity Questionnaire for Young Twins: A Research Note. Behavior Genetics, 21, 257–269. doi: 10.1007/BF01065819

Gottschling, J. (2017). *Documentation TwinLife Data: Cognitive Abilities* (TwinLife Technical Report Series, 02). Bielefeld: Project TwinLife "Genetic and social causes of life chances" (Bielefeld University / Saarland University).

Gringras, P., & Chen, W. (2001). Mechanisms for differences in monozygous twins. Early Human Development, 64, 105–117. doi: 10.1016/S0378-3782(01)00171-2

Hahn, E., Gottschling, J., Bleidorn, W., Kandler, C., Spengler, M., Kornadt, A. E., Schulz, W., Schunck, R., Baier, T., Krell, K., Lang, V., Lenau, F., Peters, A.-L., Diewald, M., Riemann, R., & Spinath, F. M. (2016). What Drives the Development of Social Inequality Over the Life Course? The German TwinLife Study. *Twin Research and Human Genetics*, *19*, 659–672. doi: 10.1017/thg.2016.76

Jackson, R. W., Snieder, H., Davis, H., & Treiber, F. A. (2001). Determination of Twin Zygosity: A Comparison of DNA with Various Questionnaire Indices. Twin Research and Human Genetics, 4, 12–18. doi: 10.1375/twin.4.1.12

Kandler, C., Riemann, R., Spinath, F. M., Bleidorn, W., Thiel, W., & Angleitner, A. (2013). The Bielefeld Longitudinal Study of Adult Twins (BiLSAT). Twin Research and Human Genetics: The Official Journal of the International Society for Twin Studies, 16, 167–172. doi: 10.1017/thg.2012.67

Lenau, F., & Hahn, E. (2017). *Documentation TwinLife Data: Zygosity* (TwinLife Technical Report Series, 01). Bielefeld: Project TwinLife "Genetic and social causes of life chances" (Bielefeld University / Saarland University).

Lichtenstein, P., De Faire, U., Floderus, B., Svartengren, M., Svedberg, P., & Pedersen, N. L. (2002). The Swedish Twin Registry: a unique resource for clinical, epidemiological and genetic studies. Journal of Internal Medicine, 252, 184–205. doi: 10.1046/j.1365-2796.2002.01032.x

Martin, N. G., & Martin, P. G. (1975). The inheritance of scholastic abilities in a sample of twins I. Ascertainment of the sample and diagnosis of zygosity. Annals of Human Genetics, 39, 213–218. doi: 10.1111/j.1469-1809.1975.tb00124.x

McGue, M., Osler, M., & Christensen, K. (2010). Causal Inference and Observational Research: The Utility of Twins. Perspectives on Psychological Science, 5, 546–556. doi: 10.1177/1745691610383511

Oniszczenko, W., Angleitner, A., Strelau, J., & Angert, T. (1993). The Questionnaire of Twins' Physical Resemblance, University of Warsaw, Poland, and University of Bielefeld, Germany. Unpublished manuscript.

Song, Y.-M., Lee, D.-H., Lee, M. K., Lee, K., Lee, H. J., Hong, E. J., Han, B., & Sung, J. (2010). Validity of the Zygosity Questionnaire and Characteristics of Zygosity-Misdiagnosed Twin Pairs in the Healthy Twin Study of Korea. Twin Research and Human Genetics, 13, 223–230. doi: 10.1375/twin.13.3.223

Spitz, E., Moutier, R., Reed, T., Busnel, M. C., Marchaland, C., Roubertoux, P. L., & Carlier, M. (1996). Comparative diagnoses of twin zygosity by SSLP variant analysis, questionnaire, and dermatoglyphic analysis. Behavior Genetics, 26, 55–63. doi: 10.1007/BF02361159

van Dongen, J., Slagboom, P. E., Draisma, H. H. M., Martin, N. G., & Boomsma, D. I. (2012). The continuing value of twin studies in the omics era. Nature Reviews. Genetics, 13, 9, 640–653. doi: 10.1038/nrg3243

Visscher, P. M., Macgregor, S., Benyamin, B., Zhu, G., Gordon, S., Medland, S., Hill, W. G., Hottenga, J.-J., Willemsen, G., Boomsma, D. I., Liu, Y.-Z., Deng, H.-W., Montgomery, G. W., & Martin, N. G. (2007). Genome Partitioning of Genetic Variation for Height from 11,214 Sibling Pairs. The American Journal of Human Genetics, 81, 1104–1110. doi: 10.1086/522934

Weiß, R. H. (2006). *Grundintelligenztest Skala 2: CFT 20-R*. Göttingen: Hogrefe.

Weiß, R. H., & Osterland, J. (2012). *Grundintelligenztest Skala 1 - Revision: CFT 1-R*. Göttingen: Hogrefe.

Xian, H., Scherrer, J. F., Eisen, S. A., True, W. R., Heath, A. C., Goldberg, J., Lyons, M. J., & Tsuang, M. T. (2000). Self-reported zygosity and the equal-environments assumption for psychiatric disorders in the Vietnam Era Twin Registry. *Behavior Genetics*, *30*, 303–310. doi: 10.1023/A:1026549417364