Simultaneous analysis of handedness data from 35 samples of twins (with a combined sample size of 21,127 twin pairs) found a small but significant additive genetic effect accounting for 25.47% of the variance (95% confidence interval [CI] 15.69–29.51%). No common environmental influences were detected (C = 0.00; 95% CI 0.00–7.67%), with the majority of the variance, 74.53%, explained by factors unique to the individual (95% CI 70.49–78.67%). No significant heterogeneity was observed within studies that used similar methods to assess handedness, or across studies that used different methods. At an individual level the majority of studies had insufficient power to reject a purely unique environmental model due to insufficient power to detect familial aggregation. This lack of power is seldom mentioned within studies, and has contributed to the misconception that twin studies of handedness are not informative.

The etiology and neurological implications of left-handedness have been the subject of debate for over a century. The first twin study comparing the handedness of monozygotic (MZ) and dizygotic (DZ) twins was published 81 years ago (Siemens, 1924); since this time there have been 34 published studies on 35 samples. However, there has been little consistency of results. The area has become one of the most controversial in laterality with the proposal of special twin effects, (such as differential right-shifts effects or mirror imaging effects; Annett, 2002; Newman, 1928) and questions raised regarding the suitability of the twin method for studying handedness and laterality in general (Nagylaki & Levy, 1973).

Comparison of results across studies is complicated by differences in methodology. Handedness can be assessed as either hand preference or hand skill; two separate but related traits. Hand preference is typically a directional measure. In its simplest form it is assessed by asking ‘are you left or right handed?’, or ‘which hand do you write with?’ In its more complex form, hand preference is assessed by asking (or asking the participant to demonstrate) which hand is preferred for a range of items. Handedness questionnaires range in length from four (Coren, 1993) to 55 items (Healey et al., 1986), the number of response choices ranging from two to five, with the responses resulting either in a handedness score/quotient or a grouping classification. The diverse methods of assessing hand preference and the number of questionnaires available within the literature in part reflects the lack of a gold standard and conceptual differences regarding the evolution and nature of handedness.

In contrast, hand skill is typically a quantitative measure of the degree of motor dominance of one hand over the other. Common tests of hand skill involve measuring the time taken to move a series of pegs on a specially designed board (Annett, 1985), or counting the number of circles that can be dotted with a pen in 20 seconds (Tapley & Bryden, 1985). Typically, the task(s) is performed with each hand and a dominance or lambda score is calculated expressing the direction and degree of dominance. The 35 published samples assessing handedness in twins are summarized in Table 1. Of the 35, one has assessed hand skill, 18 have assessed self classification or hand preference for less than three activities, 10 used questionnaires which considered hand preference for three or more activities, and two assessed handedness through observation or parental report. In the remaining four, the method used to assess handedness was not described.

Sociocultural differences across studies further complicate the comparison of results, with social attitudes towards left-handedness differing markedly both across cultures and within cultures over time. In western cultures at the beginning of the last century, left-handedness was considered highly undesirable and left-handed students were often made to write with their right hands. Attitudes towards left-handedness softened midway through the century and the pressure to be right-handed decreased dramatically, with the prevalence of left-handedness rising from
While these methodological issues make it difficult to draw clear conclusions from the literature, the overarching problem has been the lack of statistical power associated with small sample sizes. Given that the majority of twin studies have used a binary handedness classification (either left vs. right or right vs. nonright) the issue of sample size is nontrivial when trying to determine the confidence that can be placed in results. For example, given a trait with a 10% prevalence (which is typical of left-handedness) where 30% of the variance is accounted for by an additive genetic effect, about 1000 pairs of twins would be required to reject a purely unique environmental model with 80% power (Neale et al., 1994). With few exceptions, sample sizes have not been adequate to detect genetic or environmental effects that account for less than 50% of the total phenotypic variance with 80% power. While the sample size required to detect small to medium genetic effects may be beyond the resources of any one research group, such sample sizes can be reached through collaboration (Medland et al., 2003) or meta-analysis (McManus, 1980; Sicotte et al., 1999).

In 1999 Sicotte and colleagues presented an excellent meta-analysis of the 28 twin studies of handedness published at that time. Since then, data on seven additional samples have been published. Sicotte et al. (1999) addressed the issues of mirror imaging, twin sibling differences and the evidence for genetic effects. They found no evidence of mirror imaging, an increase in left-handedness in twins compared to singletons and higher concordance among MZ than DZ twin pairs, concluding that handedness is subject to genetic influences. The aim of the present article is not to redress these questions but rather to estimate the magnitude and nature of this genetic effect through joint analysis. Biometrical genetic models were applied to the published contingency table data from each study and the proportions of variance accounted for by additive genetic, unique environmental and common environmental or dominant genetic influences were estimated for each study individually, and then for all 35 samples simultaneously (such modeling techniques had been applied in only four of the studies reviewed: Bishop, 2001; Bishop, in press; Medland et al., 2003; Neale, 1988).
Orlebeke (2002), the larger of the two samples was analyzed. Details of the studies included in the analysis are given in Table 1.

While the majority of the studies assessed hand preference as a binary construct, some utilized multicategory ordinal variables or continuous variables. In the review by Sicotte et al. (1999) the data from these studies were reduced to binary variables. For the purposes of consistency and to avoid problems associated with the nonnormal nature of hand preference data (which is J-shaped) the reclassifications made by Sicotte et al. (1999) have been used in these analyses. In addition, the data of Bishop (2001) were reduced to a binary classification.

Table 1

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Handedness criteria</th>
<th>MZ</th>
<th>DZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Siemens, 1924)</td>
<td>Not stated</td>
<td>RR 41 9 16 13 2</td>
<td>LR/RL 9 4 11 6 0</td>
</tr>
<tr>
<td>(Dahlberg, 1926)</td>
<td>Hand used to cut and throw</td>
<td>RR 53 12 111 16 1</td>
<td>LR/RL 15 6 28 10 0</td>
</tr>
<tr>
<td>(von Verschuer, 1927)</td>
<td>Not stated</td>
<td>RR 58 15 6 28 10 0</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Hirsh, 1930)</td>
<td>Self-classification</td>
<td>RR 25 18 51 7 0</td>
<td>LR/RL 18 6 28 10 0</td>
</tr>
<tr>
<td>(Wilson &amp; Jones, 1932)</td>
<td>Throwing hand</td>
<td>RR 56 13 97 24 2</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Stocks, 1933)</td>
<td>Writing hand</td>
<td>RR 35 6 17 6 0</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Komai &amp; Fukushima, 1934)</td>
<td>Hand used for brush writing</td>
<td>RR 112 6 0 60 1 1</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Newman et al., 1937)</td>
<td>Hand skill — tapping task</td>
<td>RR 30 17 3 39 11 0</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Rife, 1940)</td>
<td>Hand used for 9 tasks</td>
<td>Any L responses = LH</td>
<td>RR 176 41 6 104 39 3</td>
</tr>
<tr>
<td>(Thyss, 1946)</td>
<td>Not stated</td>
<td>RR 72 24 7 60 24 2</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Rife, 1950)</td>
<td>As in Rife, 1940</td>
<td>RR 261 76 6 164 46 2</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Dechaume, 1957)</td>
<td>Not stated</td>
<td>RR 19 12 2 21 11 1</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Zazzo, 1960)</td>
<td>Self-classification</td>
<td>RR 199 51 9 264 69 2</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Koch, 1966)</td>
<td>Observation and parent report</td>
<td>RR 28 3 4 45 6 4</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Carter-Saltzman et al., 1976)</td>
<td>Writing hand</td>
<td>RR 132 46 9 115 54 7</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Loehlin &amp; Nichols, 1976)</td>
<td>Self-classification</td>
<td>RR 380 123 11 261 70 2</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Springer &amp; Searleman, 1978)</td>
<td>Writing hand</td>
<td>RR 53 19 3 35 9 3</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Hay &amp; Howie, 1980)</td>
<td>Handing criteria</td>
<td>RR 9 7 0 10 3 0</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Osborne, 1980)</td>
<td>Self-classification</td>
<td>RR 76 27 4 90 40 0</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Boklage, 1981)</td>
<td>Self-classification</td>
<td>RR 145 45 24 132 69 13</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Shimizu &amp; Endo, 1983)</td>
<td>Questionnaire</td>
<td>RR 57 4 1 41 7 0</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Forrai &amp; Bankovi, 1983)</td>
<td>Self-classification</td>
<td>RR 78 16 2 44 21 3</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Tambs et al., 1987)</td>
<td>Writing hand</td>
<td>RR 175 21 1 171 32 0</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(NCD as cited by McManus, 1985)</td>
<td>Writing hand</td>
<td>RR 32 9 2 66 18 4</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Neale, 1988)</td>
<td>Self-classification</td>
<td>RR 655 158 23 626 183 23</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Derom et al., 1996)</td>
<td>Self-classification</td>
<td>RR 249 86 17 276 108 23</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Carlier et al., 1996)</td>
<td>Writing hand</td>
<td>RR 48 6 1 15 9 0</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Orlebeke et al., 1996)</td>
<td>Self-classification</td>
<td>RR 475 122 25 764 255 22</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Ross et al., 1999)</td>
<td>Hand preference for 5 tasks: write, draw, throw, scissors, toothbrush</td>
<td>RR 923 214 21 805 155 13</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Reiss et al., 1999)</td>
<td>Hand preference for 12 tasks</td>
<td>RR 28 5 0 58 9 0</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Basso et al., 2000)</td>
<td>Self-classification</td>
<td>RR 1049 336 19 1762 247 16</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Bishop, 2001) Manchester sample</td>
<td>Number of items on a questionnaire completed with the right hand</td>
<td>RR 35 14 1 65 14 7</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Bishop, 2001) Cambridge sample</td>
<td>Writing hand</td>
<td>RR 36 9 0 33 10 1</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Medland et al., 2003)</td>
<td>Throwing hand</td>
<td>RR 1894 534 89 2050 632 80</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
<tr>
<td>(Bishop, in press)</td>
<td>Writing hand</td>
<td>RR 67 27 6 62 24 8</td>
<td>LR/RL 16 6 28 10 0</td>
</tr>
</tbody>
</table>

Notes: † Pairwise data were obtained by contacting the authors. The Basso et al. (2000) data presented here include data on all available twin pairs born before 1910. The pairwise data from Bishop (2001, Cambridge sample; 2005) reported here are only for the writing hand item, the original published data reports a laterality quotient. For the Bishop (2001) Manchester sample, participants who completed less than 6 items with the right hand are coded as left-handers. The original published data reports the full distribution.
In order to estimate the proportion of phenotypic variance arising from additive (A) and nonadditive (dominant D) genetic sources and shared (C) and non-shared (E) environmental influences, the data from previous studies were entered into a joint analysis. Because our twin data are assumed to come from MZ and DZ twins raised together, the effects of C and D are confounded and cannot be estimated together. Both ACE and ADE models in which the prevalence of left-handedness of MZ and DZ twins were allowed to differ, between and within studies, were fit to the data.

The amount of variance explained by A, C or D, and E in each sample were calculated from $2 \times 2$ contingency tables for MZ and DZ twins using Mx (1.54). These analyses were conducted within the framework of the multifactorial threshold model which posits a continuous normally distributed liability for laterality on which thresholds are imposed that define the prevalence of different definitions of handedness. The procedure is readily extended to multiple groups so that hypotheses about equality of variance components between studies can be tested. Variance components were estimated from each study individually. The variance components were then equated across studies to test for heterogeneity, and pooled variance components were estimated for all studies jointly.

### Results

#### Individual Studies

The proportion of variance accounted for by A, C and E from the ACE model within each study, the 95% confidence intervals surrounding these estimates, and the chi-square fit of the model are summarized in Figure 1. An ACE model provided a good fit to the data from all studies, except that of Osborne (1980; $= 7.95, p = .02$). There were no obvious differences between this and other studies, and the lack of fit reflects the large negative correlation observed in the DZ twins which is incompatible with both genetic and environmental models ($r = -.98; 95\% CI = -1.00, -0.20$). This pattern of data may reflect the effects of co-twin competition within this sample, or selection bias given the small sample size. The proportion of variance accounted for by A, D and E, 95% confidence intervals, and the chi-square fit of the model are summarized in Figure 2. The ADE model provided a poor fit for three studies; Komai and Fukuoka (1934; $= 6.76, p = .03$), Osborne (1980; $= 7.9, p = .02$), and Bishop (2001, Manchester sample; $= 8.31, p = .02$).

As shown in the forest plots the variance components estimates varied widely, with the lower confidence intervals on A and C seldom higher than zero in the ACE model. While for the ADE model the

---

**Figure 1**

Sample size, handedness criteria, model fit (*p < .05, 2df), and standardized estimates of additive genetic, common environmental and unique environmental variance components (with 95% confidence intervals given by the horizontal bars) by study.

The relative size of the sample is indicated by the size of the data point. For the estimates derived from the joint analysis, the 95% confidence intervals are given by the width of the polygons.
lower confidence intervals were zero for A and D across studies. Individually less than half of the studies could reject a purely environmental model. There was no clear effect of year of publication which may reflect the fact that year of publication is not always a good proxy measure of participant age/birth cohort (i.e., year of birth ranged from 1896 to 2000 in Medland et al., 2003, and 1870 to 1910 in extended sample of Basso et al., 2000). Unsurprisingly, sample size appears an important factor in detecting familial aggregation (an upper confidence interval for E < 1). Sixty-six per cent of studies with 500 or more participants and all of those with a sample greater than 1000 pairs found familial aggregation (using an ACE model), as compared to 19% of those with less than 500 pairs. However, few studies had sufficient power to determine whether the familial aggregation was genetic or environmental in nature. In the ACE model for all 35 studies the 95% CI surrounding C encompassed zero.

Joint Analysis

ACE Model: The data from each study were entered into a joint analysis (in which each group was modeled separately resulting in 70 data groups). The variance components were first equated across studies that used similar methods of assessing handedness: Self-classification, handedness based on less than three activities, handedness for longer questionnaires, and other methods. As shown in Table 2 equating estimates in this manner did not result in any significant differences. Similarly variance components could be equated across all studies regardless of the method used to assess handedness. This resulted in a change in ~2 log-likelihood of 70.64 for 68 degrees of freedom (102 equated parameters — 34 constraints; p = .39).

Across studies additive genetic factors accounted for 25.47% of the variance (95% CI 15.69–29.51%), no significant common environmental effect was found (C = 0.00; 95% CI 0.00–7.67%), with the largest proportion of variance, 74.53%, explained by unique environmental effect (95% CI 70.49–78.67%). To examine the influence of the largest sample on these results the analysis was rerun excluding the data of Medland et al. (2003), although the confidence intervals increased there was no substantial difference in the results. The analyses were also rerun excluding the data from Osborne (1980) to determine the influence of this sample (which was inconsistent with an ACE model) on the simultaneous analysis, once again though the confidence intervals shifted slightly there was little change in the variance component estimates.
ADE Model: The data from each study were entered into a joint analysis using the same procedure as described for the ACE Model. As shown in Table 2 equating estimates did not result in any significant differences ($\Delta-2LL$ 66.60 for 68 degrees of freedom; 102 equated parameters — 34 constraints; $p = .53$). Across studies additive genetic factors accounted for 19.7% of the variance (95% CI 0.50–29.3%), no significant dominant genetic effect was found (D = 6.4; 95% CI 0.00–27.1%), with the largest proportion of variance, 73.9%, explained by unique environmental effect (95% CI 69.50–78.40%). Rerunning the analyses excluding the data of Medland et al. (2003), led to no substantial differences in the results. Similarly excluding the data of Osborne (1980), Komai and Fukuoka (1934), and Bishop (2001, Manchester sample) did not alter the results. Dominant genetic effects could be dropped from the model without significant loss of fit ($\Delta-2LL$ 0.37 for 1 degree of freedom). Thus, the most parsimonious model was one in which the familial aggregation for hand preference was explained by additive genetic influences.

Discussion

Joint analysis of handedness in 35 samples of MZ and DZ twins has shown that the data are consistent with an additive genetic model of familial aggregation, in which around 25% of the variation in liability to left-handedness is explained by additive genetic influences. Somewhat surprisingly, no significant heterogeneity was observed between studies, suggesting that any cohort and or cultural differences in genetic contribution to liability between studies were minimal.

The nature of the data (contingency tables from published studies) did not allow the prevalence of left-handedness to be corrected for known covariates such as sex, cohort (year of birth), and birthweight. However, allowing the prevalence to vary freely across studies and across zygosity groups within studies would have allowed for covariate differences at a
group level. Similarly, the structure of the data did not allow an investigation of the effects of covariates on the proportions of variance accounted for by genetic and environmental effects, thereby limiting the scope of the current analyses. In the absence of Gene \times Environment interaction, the multifactorial threshold model predicts homogeneity of variance components across studies with different observed prevalence of left-handedness (i.e., differing thresholds).

One of the main conclusions that may be drawn from this analysis is that, at an individual level, it is clear that many studies have not had sufficient power to detect familial aggregation. This lack of power is seldom mentioned within studies, and should not be considered common knowledge. This unacknowledged limitation has contributed to the belief that twin studies of handedness are not informative.

A number of competing genetic models have been proposed in the literature (Annett, 1985; Crow, 2002; Klar, 1999; McManus, 1985) and these have been applied to some of the data described here (e.g., Annett, 2002), with the results generally showing an acceptable fit of one or more of the models. While the polygenic ACE model utilized here provided a good fit to all but one of the studies, the lack of power associated with many studies limits the conclusions that can be made regarding the polygenic nature of the genetic influences on handedness. Linkage analysis of hand skill by Francks et al., (2003, 2002) has found a significant parent of origin (maternally imprinted) linkage to the chromosomal region 2p12–q11, with the second most significant linkage on chromosome 17. Although the genetic correlation between hand skill and hand preference has not been thoroughly investigated in unselected samples, the results of linkage analysis for hand skill are inconsistent with the models of Annett, McManus, Klar and Crow, and suggest that hand skill, and by association hand preference, are complex traits that may be influenced by more than one gene.

Acknowledgments

The authors wish to thank Professor Dorothy Bishop, and Professor Kaare Christensen, Dr Olga Basso, and Dr Inge Petersen for their assistance in providing data for these analyses.

Sarah Medland was supported in part by a Thenie Baddams Grant from the Australian Federation of University Women — SA Inc.

References


