



# Returns to Education: What Do Twin Studies Control?

Kevin C. Stanek, William G. Iacono and Matt McGue  
*Department of Psychology, University of Minnesota, United States of America*

The current article examines the assumptions of the MZ co-twin control method that has been applied in attempts to gain more accurate estimates of the returns to education by naturally ‘controlling’ for individual differences on which MZ twins are matched. The current study examined 1738 MZ and 926 DZ twins from Minnesota, including 133 pairs of MZ and 101 pairs of DZ twins discordant for university attendance. They were assessed prospectively on personality, intelligence, GPA, and academic motivation; a subset also has reported income at age 29. MZ twins discordant for university attendance differed significantly and prospectively on verbal IQ, personality traits, and GPA. While MZ co-twin control studies can provide more accurate estimates of the returns to education than analyses of single individuals, these studies do not entirely obviate the need to control for differences between university students and non-students that predate university attendance and might account for income differentials and even non-monetary outcomes.

■ **Keywords:** co-twin control method, discordant twins, returns to education, income

The most straightforward method of assessing the value of education has been to examine the differences in income of those with varying amounts of education. One obvious problem arising from this method is that individuals who differ in amounts of education consumed and income earned may also systematically differ on other dimensions, such as cognitive ability or personality, confounding the comparison. Ideally, researchers would like to know what an individual’s income would have been if they had pursued additional education and, in a separate but identical reality, what that individual’s income would have been had they not pursued additional education. Such perfect counterfactuals do not exist, so scientists have investigated proxies, such as twins using the MZ co-twin control method.

As a result of their matching genetic background and rearing environment, MZ twins are often used as counterfactuals for one another in experimental designs such as the MZ co-twin control (McGue et al., 2010). Numerous twin studies in several nations have produced estimates of the economic benefits of education, and all attempt to reduce the error variance inherent in such estimates by reducing biasing individual differences (Ashenfelter & Krueger, 1994; Isacson, 2004; Miller et al., 2006). In general, ordinary least squares estimates of the returns to education based on unrelated individuals are around 13% (i.e., an individual’s annual income can be expected to rise

13% with each additional year of schooling, controlling for gender and year of birth; Leigh & Ryan, 2008). When such individual-level estimates are also adjusted for marital status and full-time worker status, returns to education fall to around 10% in Western countries (Leigh & Ryan, 2008). Further, when MZ twins who are discordant for level of education are used to estimate the returns to education by comparing the income of the twin with more education to the income of the twin with less education, researchers often report estimates around 5% (Miller et al., 2006). Estimates of the returns to education vary based on the cohort examined, level of country development, demographic variables included in the model, types of corrections made to adjust for errors in measurement, and numerous other variables, but in all, the estimates often fall in the 3–15% range (Ashenfelter et al., 1999). Addressing previous concerns by Griliches (1979) and later echoed by Neumark (1999) and others, the current research investigates the tenability of the assumptions that

RECEIVED 21 October, 2010 ; ACCEPTED 07 October, 2011.

ADDRESS FOR CORRESPONDENCE: Kevin C. Stanek, Department of Psychology/Elliott Hall, University of Minnesota, Minneapolis, MN 55455, USA. E-mail: stane040@umn.edu

underlie the co-twin control method and may contribute to these differing estimates. Specifically, we examine if twins are similar enough before university to attribute post-university income differences to differences in levels of education. The answer to this query also provides information as to how well the co-twin control method reduces error variance due to individual differences.

The twin difference, or MZ co-twin control, method was logically applied to the problem of estimating the returns to education for several reasons. Primarily, it became apparent that ability, environment, personality, and other factors were biasing estimates based on comparisons of unrelated individuals (Card, 2001). MZ twins offer an appealing solution to these confounding issues as MZ twins are assumed to have similar environments and identical genetic material. Given these similarities, any resulting difference in income is more likely attributable to differences in amount of education between the twins. Such a string of logic holds as long as prior to the decision of attending university the twins are matched on variables that are relevant to later income. However, such assumptions are idealistic. More broadly, even MZ twins are likely to differ to some degree in personality, behavior, and abilities and these differences are almost certainly related to later outcomes. If it can be established that MZ twins are significantly different on dimensions related to university attendance that are also related to later life outcomes, then the MZ co-twin control method will not completely account for relevant confounders in estimating the returns to education, both financial and social.

We hypothesize that MZ twins differ significantly on several important dimensions before university age, that these differences are related to university attendance, and that these pre-existing differences are also related to later income levels and could thus be biasing MZ co-twin control method estimates of the returns to education. Furthermore, we predict that the MZ twin who attends university will have higher income compared to their non-university attending twin. Within the final prediction, we recognize that MZ twins are highly similar in intelligence ( $r \sim .80$ ), so we expect the MZ co-twin control method to work well in that regard, although it may not perform as well regarding factors such as personality, which MZ twins are less similar on ( $r \sim .50$ ) (McGue & Bouchard, 1998).

If the MZ co-twin control method fully accounted for confounders when estimating the returns to education, we would expect to see no pre-existing differences within a twin pair on variables that relate to later income in twins before university attendance. The present study investigated a large, longitudinal sample of twins and asked two main questions: 1) To what extent are twins who are discordant for university attendance matched on variables that predict going to university, and 2) If they are not matched, are the differentiating factors also related to later income?

## Methods

### Sample

Data from 11-year-old and 17-year-old cohorts of the Minnesota Twin Family Study (MTFS) were used for the current analyses. For additional information on the design of the MTFS, including participation and attrition analyses please see other articles that more precisely report the nature of the study (e.g., Iacono et al., 1999; Iacono & McGue, 2002). The current sample is largely Caucasian, 52.2% female, and consists of the 869 MZ and 463 DZ pairs from the MTFS for which we have information on university attendance at age 20 (95% of the total MTFS sample). The sample includes twins from both the younger (initial age of assessment at age 11) and older (initial age of assessment at age 17) cohorts of the MTFS, and is based on data collected at the age 17, 20 and 29 assessments of the two cohorts. While the age-29 assessment is complete for the older MTFS cohort, it is still underway for the younger cohort (see Table 1). Consequently, only a portion of the sample has information for the age-29 outcome, specifically income, and our analyses of these data should be considered preliminary. Discordant pairs were defined by one twin reporting having attended a 4-year university and the other indicating they had not at the age-20 assessment. There were 133 MZ and 101 DZ university-discordant pairs. Among the university-discordant pairs, both members of only 48 MZ and 32 DZ pairs completed the age-29 assessment and were working full time, thus rendering examination of the income differences within these discordant pairs more illustrative than definitive.

### Measures

The current study gathered self-report data on personality, intelligence, academic achievement (GPA), academic motivation, educational attainment, and annual income. Personality, intelligence, grade point average, and academic motivation were used to predict educational attainment/university attendance and were obtained at the age-17 assessment, when the twins were in their senior year of high school. Educational attainment information was gathered during age-20, and verified during age-25, interviews, while income information was gathered in age-29 interviews.

**TABLE 1**

Description of the Sample

	Birth years	N (%) Attending university at age 20	N (%) completing Age-29 assessment and providing income data
Younger cohort	1975–1979	738 (50.9%)	104 (7.2%)
Older cohort	1972–1977	549 (44.2%)	1028 (82.7%)

**Multidimensional Personality Questionnaire.**

Personality traits were assessed using the Multidimensional Personality Questionnaire (MPQ; Tellegen & Waller, 2008). The MPQ consists of 198 items answered on a four-point scale (ranging from *Definitely True* to *Definitely False*), which make up 11 primary scales organized into three higher-order scales, Positive Emotionality, Negative Emotionality, and Constraint. The MPQ dimensions have been shown to have high test-retest reliability (30 day median value of 0.89), high internal consistency ( $\alpha > .75$  for all dimensions), and high external correlations (e.g., observer ratings and other personality inventories) (Tellegen & Waller, 2008). In the current study, we used the method described by Jockin et al., (1996) to compute a personality composite that maximized predictive power with outcome (here university attendance) while minimizing the number of variables used (here personality scales). Specifically, in a logistic regression we identified those MPQ primary scales that significantly and independently predicted university attendance. These scales were Achievement, Aggression, Alienation, and Control. A personality composite, reflecting the combined contribution of personality factors to the prediction of university attendance, was then formed by taking a weighted linear composite of these four scales with weights given by the corresponding logistic regression coefficients.

**Wechsler Adult Intelligence Scale.** Intelligence was assessed using a verbal score made up of the Vocabulary and Information subscales of the Wechsler Adult Intelligence Scale-Revised, shown by previous research to be the two highest loading scales on a higher order Verbal factor (Sattler, 1974). Verbal intelligence was chosen because it was more predictive of university attendance than the Performance or Full score, and neither of the latter added to the prediction of university attendance after verbal intelligence had been accounted for.

**Academic Motivation.** Academic motivation was assessed using a self-report Childhood Academic History Questionnaire. The Childhood Academic History Questionnaire aims to measure psychological engagement in school, specifically: interest, involvement, and effort at age-17. Typical items include ratings of how true (*Definitely true of me, Probably true of me, Probably false of me, Definitely false of me*) an attitude such as, 'Turn in my homework on time' or 'Easily distracted in class' are for the individual. The estimated internal consistency reliability of this measure is 0.74 for girls and 0.85 for boys. Previous research has used this variable in a similar population to demonstrate a significant relationship with school grades (Johnson et al., 2006).

**Grade Point Average (GPA).** Since US high schools differ in grading standards, systems, and procedures, report

cards were not used to obtain this data. Instead, participants and one of their parents rated each academic subject (i.e., language arts, math, social studies, and science) in terms of the twin's typical standing relative to their peers. A rating of 4 indicated grades that were 'very much above average, one of the top in the class' (i.e., A), a rating of 3 indicated grades that were 'above average' (i.e., B), a rating of 2 indicated grades that were 'average' (i.e., C), and a rating of 1 indicated that grades were 'below average' (i.e., D), and a rating of 0 indicated that grades were 'very much below average, failing' (i.e., F). These ratings were averaged across academic subjects to create the GPA variable. To verify the validity of this method, 67 twins were randomly selected to have their rated GPA verified by a school transcript. Actual GPA and self-rated GPA correlated 0.89. Individuals who were no longer in high school at their age-17 assessment rated their GPA for the last year they attended high school (Hicks et al., 2008).

**University attendance.** Educational attainment was obtained during the age-20 interview (verified at age-25 for those who said that they had attended) that asked participants, 'Have you ever attended a 4-year college or university?' Only pairs where both twins provided all necessary data could qualify as 'discordant'.

**Annual income and full time status.** Income was reported by participants at age-29: 'What is your annual income before taxes?' Age-29 data was available for 47% (825) of the MZ twins and 47% (436) of the DZ twins, as this assessment is still being completed by the younger MTFs cohort. Reported incomes were deciled to reduce the influence of outliers. Full time employment status was determined by direct question: 'In the past 12 months, how many months were you employed *full time*?', where only individuals answering '12' were counted as full-time workers. When only full-time workers are examined, income data was available for 596 MZ and 317 DZ twins.

**Zygoty.** Zygoty was primarily determined through questionnaires, investigator estimates, and anthropometric measurements. If these three methods did not agree, DNA was collected from blood samples and zygoty determined using genetic markers.

**Method of Analysis**

Initially, to identify factors related to university attendance; verbal IQ, personality, GPA, and academic motivation were correlated with university attendance at the individual-level. These predictor variables were then correlated with adult income. After demonstrating these relationships, twin correlations were assessed for all variables and pairs discordant for university attendance were compared. Mean differences on the standardized predictor variables were analyzed to see if differences in educational attainment, and thus potentially later income, could be

partially attributable to pre-existing differences between the twins.

To facilitate interpretation of results, quantitative variables were standardized when appropriate for the analyses. Predictor interactions with sex were examined but found to not reach levels of statistical significance. Consequently, all findings reported here are for the combined sample of males and females.

**MZ Co-Twin Control**

The MZ co-twin control design capitalizes on the genetic and environmental similarity of MZ twins. The basic premise is that two individuals randomly selected from the population will vary greatly in a near-infinite number of ways; in comparison, MZ twins have been shown to exhibit high degrees of similarity across a variety of measures. Studying individuals who are similar reduces the number of possible explanations for later differences. In the case of MZ pairs, twins are matched on genotype and rearing (i.e., shared) environment. Consequently, the effects of these factors are controlled when comparing the two members of an MZ pair. If the factor on which the MZ pairs are discordant is causally related to outcome, then we would expect to see outcome differences between the two members of a pair. In the case of DZ pairs, the twins are matched perfectly on shared environment but only partially (i.e., 50%) on genotype. Consequently, the control for DZ pairs is not as exact as it is for MZ pairs. Nonetheless, the DZ pairs are informative, as failure to observe differences with DZ (and MZ) discordant pairs would implicate the importance of shared environmental confounders, while observing greater within pair effects in DZ than MZ twin pairs would implicate the importance of genetic factors. In either case, it is important to recognize that the co-twin control method does not provide a certain test for

causality. This is because the causal agents may be factors that led to the discordance, rather than discordance per se. In this study we sought to determine whether the factors that led to differences in twins' educational attainment may be confounding estimates of return to education even when using a co-twin control design.

**Results**

**Predictors of University Attendance at the Individual Level**

We first investigated whether each of the predictor variables was associated with university attendance in the full sample (i.e., DZ and MZ twins) by comparing mean scores of university attendees and non-attendees. All of the predictor variables (i.e., verbal intelligence, personality composite, grade point average, and academic motivation) were found to be significantly ( $p < .001$ ) related to university attendance (see Table 2).

**Twin Resemblance**

Twin correlations were calculated for the predictor variables (Pearson) as well as for university attendance (tetrachoric) and are given in Table 3. The MZ twin correlations were significantly higher ( $p < .001$ ) than the DZ twin correlations in all cases indicating genetic influence on the traits. Moderately large MZ twin correlations for all of the predictor variables also suggest that the MZ co-twin control method would likely provide partial, but not complete, control for the confounding effects of these variables on estimates of the returns to education.

**Twins Discordant for University Attendance**

In pairs where one twin attended a 4-year university by age-20 and the other did not, differences in the predictor variables were examined (see Table 2). Note that approxi-

**TABLE 2**  
Predictors' Relationships with University Attendance

Age 17 predictor	Individual level					Discordant MZ pairs				Discordant DZ pairs					
	University		Non-University			N pairs	University		Non-University		N pairs	University		Non-University	
	N	$\bar{x}$ (SD)	N	$\bar{x}$ (SD)	Stdized Mean Diff (SE)*		$\bar{x}$ (SD)	$\bar{x}$ (SD)	Stdized Mean Diff (SE)*	$\bar{x}$ (SD)		$\bar{x}$ (SD)	Stdized Mean Diff (SE)*		
Grade point average	1199	3.40 (0.55)	1029	2.71 (0.73)	0.934* (0.047)	118	3.17 (0.58)	3.05 (0.60)	0.160† (0.073)	92	3.33 (0.51)	2.72 (0.67)	0.839* (0.102)		
Verbal IQ	1224	101.43 (13.48)	1097	90.31 (11.72)	0.802* (0.044)	122	98.20 (13.25)	95.75 (12.36)	0.177† (0.052)	95	98.96 (12.17)	94.04 (11.38)	0.345* (0.090)		
Personality	1213	100.03 (65.78)	1068	57.21 (70.49)	0.581* (0.045)	111	78.71 (64.45)	53.42 (71.61)	0.355* (0.090)	93	97.54 (62.62)	48.58 (72.06)	0.688* (0.108)		
Academic motivation	1176	20.36 (2.58)	1040	18.79 (3.40)	0.485* (0.051)	117	19.86 (2.78)	19.65 (3.00)	0.069 (0.100)	94	20.37 (2.52)	18.81 (3.49)	0.504* (0.117)		

Note: \* Standardized mean differences were significant at the  $p < .001$  level; † Standardized mean differences were significant at the  $p < .01$  level; Grade point average was scaled from 0-4 and was reported by the twins and caregivers, with a 4 being 'one of the top in the class'; Verbal IQ was scaled as the composite of an individual's vocabulary and information subscale scores on the WAIS-R. Personality was scaled as a composite of Achievement, Alienation, Aggression, and Control, which were each canonically weighted. Academic motivation was assessed using self reports from the Childhood Academic History Questionnaire which gathers information regarding effort, interest, and involvement in school.

TABLE 3

## Twin Resemblance

Predictor	$r_{MZ}$ ( <i>n</i> = 730–784 pairs)	$r_{DZ}$ ( <i>n</i> = 380–410 pairs)	$\chi^2$	$\chi^2$ <i>P</i> value	$r_{MZ\text{ Discordant}}$ ( <i>n</i> = 111–122 pairs)	$r_{DZ\text{ Discordant}}$ ( <i>n</i> = 92–95 pairs)	$\chi^2$	$\chi^2$ <i>P</i> value
GPA	0.68	0.41	40.18 (df = 1)*	< .001	0.53	0.31	3.64 (df = 1)	.056
Verbal IQ	0.84	0.57	90.94 (df = 1)*	< .001	0.81	0.47	20.08 (df = 1)*	< .001
Personality	0.55	0.28	28.06 (df = 1)*	< .001	0.52	0.40	1.08 (df = 1)	.299
Academic motivation	0.49	0.26	18.00 (df = 1)*	< .001	0.33	0.35	0.35 (df = 1)	.555
University <sup>1</sup>	0.86	0.74	32.26 (df = 1)*	< .001				

Note: All correlations are significant at the  $p < .04$  level; \* Correlation differences significant at the  $p < .001$  level.  
University  $r_{MZ}$  and  $r_{DZ}$  computed as tetrachoric correlations.

mately 18% of pairs were discordant for university attendance. In the DZ discordant twin sample, all predictor variables were significantly associated with university attendance; standardized mean differences were moderate to large. In the MZ discordant twin sample, the personality composite had a significant effect ( $p < .001$ ), and verbal IQ and GPA also significantly differentiated the two members of the twin pair ( $p < .01$ ); the difference was non-significant for academic motivation. In the MZ discordant twin sample, standardized mean differences were generally small, indicating that the use of discordant MZ twins should largely but not entirely control for differences on these variables.

#### Relationships of Predictors of University Attendance with Later Income

Correlations between the university predictors at age-17 and deciled income at age-29 are given in Table 4. Verbal IQ, personality, and academic motivation were significantly correlated with income ( $p < .05$ ). GPA showed the strongest relationship ( $p < .001$ ), although all of the relationships were modest in magnitude. GPA, verbal IQ, and personality were also significantly related to raw income. For comparison purposes, we have also included in Table 4 the correlation between each of the predictors and university attendance (dichotomously coded). As can be seen, all variables are much stronger predictors of pursuing further education than of age-29 income.

#### Income Differences for Individuals Discordant for University Attendance and Age-29 Income

Comparison of age-29 annual salaries of individuals who did versus did not attend university is given in Table 5. At

the individual level, university attendance was associated with a 16.1% or \$6,739 average increase in annual income. We also report in Table 5 income results for discordant twins, showing that income differences were significant within discordant DZ but not within discordant MZ pairs. However, since income data were available on less than half of our sample at the time we completed these analyses, we consider these results to be preliminary and include them here to illustrate the application of co-twin control method.

## Discussion

Discussions of the true monetary returns to education have progressed for decades and been advanced by many methods. One approach that researchers have employed is the use of twin studies, aiming to more precisely estimate the value of additional years of schooling. Such studies, however, are based on relatively untested assumptions. This paper has specifically examined the assumption that MZ twins can viably serve as counterfactuals for one another, that is, that twins are similar enough on dimensions associated with attending university and later income that they can be considered controls for one another (at least before the disparate years of education). We found that MZ twins discordant for university attendance were similar on most of the factors related to university attendance and later income, suggesting that the MZ co-twin control design provides an approximate, albeit not perfect, control for relevant confounding variables.

We first demonstrated that several cognitive (i.e., verbal IQ and high school GPA) and non-cognitive factors (i.e., academic motivation and personality) significantly differentiated those who attended university from those who

TABLE 4

## Relationships Between Predictors and Outcomes at the Individual Level

Predictor	$r_{\text{University Attendance}}$	$r_{\text{Deciled Income}}$	$r_{\text{Raw Income}}$
GPA	0.476* (N = 2228)	0.115† (N = 1080)	0.099* (N = 1080)
Verbal IQ	0.401* (N = 2321)	0.077† (N = 1112)	0.081† (N = 1112)
Personality	0.300* (N = 2281)	0.084† (N = 1020)	0.077† (N = 1020)
Academic motivation	0.253* (N = 2216)	0.076† (N = 1093)	0.039 (N = 1093)

Note: \* Indicates a standardized mean difference significant at  $p \leq .001$ ; † Indicates a standardized mean difference significant at  $p < .05$ .

**TABLE 5**

Preliminary Findings Regarding University Discordance and Age 29 Income for Full Time Workers

Predictor	Individual level					MZ twins					DZ twins				
	University		Non-University		Stdized Mean Diff (SE)*	University		Non-University		p value	University		Non-University		p value
	N	$\bar{x}$ (SD)	N	$\bar{x}$ (SD)		N	$\bar{x}$ (SD)	$\bar{x}$ (SD)	Stdized Mean Diff (SE)*		N	$\bar{x}$ (SD)	$\bar{x}$ (SD)	Stdized Mean Diff (SE)*	
Deciled salary	528	5.10 (2.20)	534	4.37 (2.03)	0.343* (0.061)	48	5.06 (1.90)	4.94 (2.10)	0.059 (0.145)	0.689	32	5.63 (2.15)	4.06 (1.29)	0.732* (0.194)	.001
Stdized deciled salary	528	0.173 (1.03)	534	-0.170 (0.95)	0.343* (0.061)	48	0.155 (0.89)	0.100 (0.98)	0.059 (0.145)	0.689	32	0.367 (0.56)	-0.094 (0.48)	0.732* (0.194)	.001
Raw salary	528	48622 (28039)	534	41883 (32620)	0.223* (0.062)	48	50646 (36435)	52188 (48981)	-0.051 (0.146)	0.729	32	54594 (26513)	37375 (13341)	0.569* (0.160)	.001

Note: \*Indicates a standardized mean difference significant at  $p < .01$ .

did not at the individual level. We examined MZ twin resemblance on these factors and found it to be moderately high, but not to levels that might deserve the descriptor 'identical'. We also examined twin pairs that exhibited discordance for university attendance and found systematic differences differentiating twins even before differences in educational attainment emerged. Specifically, within discordant MZ twin pairs, verbal IQ, the personality composite, and GPA differentiated the twin attending university from the one who did not. These within-pair differences were, however, small and only verbal IQ, the personality composite, and GPA were consistently related to post-university income. As a result of these relationships, we conclude that the MZ co-twin control method corrects for some, but not all of the pre-existing differences in individuals who do and do not attend university. A comparison of the individual level analyses and MZ within-pair level analyses in Table 2 clearly demonstrates the control gained over confounding differences/variables by examining MZ twins. Differences in income that are found in this field of research are often attributed to differences in years of education and demographic variables, but they may partially be the result of pre-existing individual differences that are also associated with individuals pursuing different amounts of education. These pre-existing differences, and perhaps even university itself, may also lead to non-monetary/indirect returns to education. Such 'social' returns are important to keep in mind so that the value of education is not naïvely weighed only in terms of direct monetary outcomes. An important final note is that this paper differs from others examining returns to education by only comparing individuals who attended university at a certain age compared to those who did not.

In summary, the current paper shows that the application of the MZ co-twin control method to the problem of identifying the true returns to education is an imperfect method for controlling some variables, such as verbal intelligence, that are important for predicting years of education attained as well as later annual income.

Nonetheless, as a result of their genetic and rearing environment similarity, the MZ co-twin control method does reduce error variation in the estimates of returns to education for some factors that affect educational attainment and income. Upon examination, the standardized mean differences between predictor scores for discordant twins within a pair, especially MZ twins, are relatively small. Such small effect sizes indicate that the MZ co-twin control method is largely fulfilling its goal of controlling for pre-existing differences effects on the later income of individuals who do and do not attend university. The MZ co-twin control method is not perfect however, and it does not obviate the need to control for confounding variables such as personality.

## Acknowledgments

Supported by USPHS Grants # DA05147 and AA009367.

## References

- Ashenfelter, O., Harmon, C., & Oosterbeek, H. (1999). A review of estimates of the schooling/earnings relationship, with tests for publication bias. *Labour Economics*, 6, 453–470.
- Ashenfelter, O., & Krueger, A. (1994). Estimates of the economic return to schooling from a new sample of twins. *The American Economic Review*, 84, 1157–1173.
- Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69, 1127–1160.
- Griliches, Z. (1979). Sibling models and data in economics: Beginnings of a survey. *The Journal of Political Economy*, 87, 37–64.
- Hicks, B. M., Johnson, W., Iacono, W. G., & McGue, M. (2008). Moderating effects of personality on the genetic and environmental influences of school grades helps to explain sex differences in scholastic achievement. *European Journal of Personality*, 22, 247–268.
- Iacono, W. G., Carlson, S. R., Taylor, J., Elkins, I. J., & McGue, M. (1999). Behavioral disinhibition and the development of substance-use disorders: Findings from the Minnesota

- Twin Family Study. *Development and Psychopathology*, 11, 869–900.
- Iacono, W. G., & McGue, M. (2002). Minnesota twin family study. *Twin Research and Human Genetics*, 5, 482–487.
- Isacsson, G. (2004). Estimating the economic return to educational levels using data on twins. *Journal of Applied Econometrics*, 19, 99–119.
- Jockin, V., McGue, M., & Lykken, D. T. (1996). Personality and divorce: A genetic analysis. *Journal of Personality and Social Psychology*, 71, 288–299.
- Johnson, W., McGue, M., & Iacono, W. G. (2006). Genetic and environmental influences on academic achievement trajectories during adolescence. *Developmental Psychology*, 42, 514–531.
- Leigh, A., & Ryan, C. (2008). Estimating returns to education using different natural experiment techniques. *Economics of Education Review*, 27, 149–160.
- McGue, M., & Bouchard Jr, T. J. (1998). Genetic and environmental influences on human behavioral differences. *Annual Review of Neuroscience*, 21, 1–24.
- McGue, M., Osler, M. & Christensen, K. (2010). Causal inference and observational research: The utility of twins. *Perspectives on Psychological Science*, 5/5, 1745–6924.
- Miller, P., Mulvey, C., & Martin, N. (2006). The return to schooling: Estimates from a sample of young Australian twins. *Labour Economics*, 13, 571–587.
- Neumark, D. (1999). Biases in twin estimates of the return to schooling. *Economics of Education Review*, 18, 143–148.
- Sattler, J. M. (1974). *Assessment of children's intelligence*. Philadelphia: Saunders.
- Tellegen, A. & Waller, N. G. (2008). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire. In G. J. Boyle, G. Matthews & D. H. Saklofske (Eds.), *The SAGE Handbook of Personality Theory and Assessment: Vol. 2 Personality Measurement and Testing* (pp. 261–292). London: Sage.
-