Published in final edited form as:
Dev Sci. 2009 July ; 12(4): 626–633. doi:10.1111/j.1467-7687.2008.00801.x.

Do Different Measures Tap the Same Genetic Influences? A Multi-method Study of Activity Level in Young Twins

Kimberly J. Saudino
Boston University

Abstract

Activity level (AL) is a highly salient feature of child behavior that has been linked to developmental outcome. Twin studies of parent-rated, observer-rated and mechanically-assessed AL in childhood find that AL is genetically influenced. Few studies, however, consider whether different methods of assessing AL have a shared genetic etiology. Those that do, confound methods and situations. The present study examined whether actigraph and rater-based (parent, observer) measures of AL tap the same genetic influences in a sample of 312 2-year-old twin pairs. Methods were studied within the same situation thereby controlling for situational influences on AL. The genetic correlation between actigraph and parent-rated AL in the home was .38, indicating modest genetic overlap between the two methods. In contrast, genetic effects on actigraph and observer-rated AL in the lab correlated . 95 indicating that both lab-based measures of AL are influenced by the same genetic effects.

Activity level (AL) is a core dimension that appears in most theories of temperament. It is also a key behavioral component of hyperactivity (characterized by overactivity, impulsivity and inattentiveness). Although most research has focused on hyperactivity and child adjustment, there is a substantial body of work that shows that AL viewed as a temperament dimension or as the simple motoric component of hyperactivity is negatively associated with various facets of child health and development. For example, the link between AL in infancy and early childhood and behavioral problems has been well documented. High AL has been associated with externalizing problems such as aggression, hyperactivity, and conduct disorder in both clinical and nonclinical samples (Campbell & Breaux, 1983; Fagan, 1990; Fagot & O'Brien, 1994; Mehregany, 1991; Teglasi & MacMahon, 1990). Moreover, AL assessed during the first two years of life significantly predicts behavior problems in middle childhood (Rothbart, Derryberry, & Hershey, 2000; Sanson, Smart, Prior, & Oberklaid, 1993) suggesting that AL may be an important risk factor for the development of adjustment disorders. AL is also linked to other important developmental outcomes. Children who display high AL in infancy or early childhood have more conflict and negativity within social interactions in middle childhood (Rothbart, Derryberry, & Posner, 1994). Similarly, early AL is negatively associated with IQ and academic achievement assessed several years later (Halverson & Waldrop, 1976; Martin, Olejnik, & Gaddis, 1994). Thus, AL may interfere with the development of behaviors that are conducive to cognitive and social abilities in addition to behavior problems.

The developmental significance of AL in early childhood makes it important to understand its etiology. Twin studies consistently suggest that AL is genetically influenced. This is true when AL is assessed via parents, trained observers or mechanical motion recorders (Goldsmith, Lemery, Buss, & Campos, 1999; Saudino & Eaton, 1991, 1995; Saudino, Plomin, & DeFries, 1996). An implicit assumption in the literature is that the different measures of AL tap the same

Correspondence concerning this article should be addressed to Kimberly J. Saudino at the Psychology Department, Boston University, 64 Cummington St., Boston, MA, 02215. Email: E-mail: ksaudino@bu.edu.

genetic effects, but this assumption has rarely been questioned. Phenotypic correlations between parent ratings of AL and observational or mechanical measures are modest (Saudino, Wertz, Gagne, & Chawla, 2004; Seifer, Sameroff, Barrett, & Krafchuck, 1994), indicating that, to some extent, different methods are influenced by different factors. Thus, it is possible that the genetic factors that influence one method of assessing AL may differ from those that influence another.

The few behavioral genetic studies of AL that have explored the question of genetic overlap across methods have found evidence of unique genetic variance. For example, a multitrait-multimethod approach to infant twin data comprising three temperament dimensions (activity, emotional tone, sociability) assessed three different ways (tester rating, playroom observation, parent rating) found substantial method-specific genetic effects (Philips & Matheny, 1997). Similarly, a twin study examining the etiological overlap between actigraph-assessed AL and parent and teachers ratings of hyperactivity in middle childhood found that although there was genetic variance common to all three measures, there was also genetic variance specific to the actigraph and parent ratings (Wood, Rijsdijk, Saudino, Asherson, & Kuntsi, 2008). Genetic correlations between methods ranged from .21 between actigraphs and teacher ratings to .48 between parent and teacher ratings, indicating that less than 50% of the genetic effects overlap across methods. These studies, however, confound methods with situations. That is, the different methods of assessing AL were used in different situations (e.g., parent report in the home, observer report or actigraphs in the lab, teacher report in the school). Consequently, it is difficult to determine whether the unique genetic variance is specific to the method or to the situation in which AL was assessed. Indeed, Phillips and Matheny also refer to their finding of genetic method effects as “situation-specific” genetic effects. Moreover, situational-specific genetic variance has emerged from a study using the same measure of AL (actigraphs) across multiple situations (home and lab) (Saudino & Zapfe, in press) suggesting that the apparent method effects found in prior studies might be due to situational differences.

The present study explores whether different measures of AL tap the same genetic influences in early childhood by employing multiple measures within the *same* situation, thereby controlling for situational influences on AL. The AL of 2-year-old twins was assessed in the home and the lab. In both situations, actigraphs provided an objective measure of the physical forces that are associated with human activity. Parent ratings of temperament provided a second measure of AL in the home and observer ratings of AL were obtained in the lab. Multivariate genetic analyses of actigraph and rater-based AL measures within the same situation were used to estimate genetic and environmental covariances between measures. To the extent to which the different AL measures assess the same construct, it was predicted that what is in common between measures is, in part, heritable, even though the phenotypic correlation between the measures may be modest.

Methods

Sample

The Boston University Twin Project (BUTP) sample was recruited from birth records supplied by the Massachusetts Registry of Vital Records. Twins were selected preferentially for higher birth weight and gestational age. No twins with birth weights below 1750 grams or with gestational ages less than 34 weeks were included in the study. Twins were also excluded if they had a health problem that might affect motor activity (e.g., club foot) or had chromosomal abnormalities. The present analyses include 312 same-sex pairs of twins (144 MZ, 168 DZ), mean age 2.07 years (*SD* = .05). Ethnicity was generally representative of the Massachusetts population (85.4% Caucasian, 3.2% Black, 2% Asian, 7.3% Mixed, 2.2% Other). Socioeconomic status according to the Hollingshead Four Factor Index (1975) ranged from low to upper middle class (range=20.5-66; *M* =50.9, *SD*=14.1).

Zygosity was determined via DNA analyses using DNA obtained from cheek swab samples. In the cases where DNA was not available (*n*=3), zygosity was determined using parents' responses on physical similarity questionnaires which have been shown to be more than 95% accurate when compared to DNA markers (Price et al., 2000).

Procedure Overview

Twins were assessed within approximately 2 weeks of their second birthday. The procedure consisted of two visits, 48-hours apart, to the laboratory. At the initial visit, informed consent from parents was obtained and actigraphs were attached to each child. After attachment of the actigraphs, one twin was assessed within a standardized test situation, while the other twin was assessed within a play situation. The test situation involved administration of the Mental Scale of *Bayley Scales of Infant Development–Second Edition* (Bayley, 1993). The play situation comprised activity episodes (arc of toys, corral of balls, workbench, fidget video) from the *Laboratory Temperament Assessment Battery–Preschool Version* (Goldsmith, Reilly, Lemery, Longley, & Prescott, 1995). At the second visit, situations were reversed for each twin. Actigraphs were then removed and questionnaires and cheek scrapings collected. Twins were assessed by different testers, however, for each twin the tester was the same across the two laboratory situations.

Measures

Mechanical assessment of AL—During the two-day data collection period, activity level was assessed with the Minimitter actical (actigraph). This device is a miniature omnidirectional accelerometer designed to detect frequency and intensity of motion within the range of normal human movement (*Actical Physical Monitoring System Instruction Manual*, 2005). Actigraphs have been shown to be valid instruments for recording the activity of young children (Pate, Almeida, McIver, Pfeiffer, & Dowda, 2006). Because actigraphs collect real-time data, it is possible to extract activity counts for specific time periods (i.e., home and lab).

Each twin was randomly assigned four actigraphs, one per limb. Actigraphs were attached by means of tyvek© adhesive wristbands. Arm attachment, at the wrist, was on the dorsal aspect of the forearm proximal to the radial carpal joint. Leg attachment, at the ankles, was superior to the lateral malleoli. Parents were provided with instructions regarding the care of the actigraphs and sheets for logging the times when they were off a limb. To differentiate actigraphs within twin pairs, each actigraph was covered with white tape on which the child's name was printed. Actigraphs were color-coded and keyed to a parent record sheet indicating the color of actigraph that went on each limb. Parents were encouraged to maintain normal daily routines with their children.

The average length of time that the actigraphs were worn was 48.22 hours (*SD* = 3.22). To adjust for variations in the total time that each instrument was worn within the home and laboratory, the number of activity units was converted to a rate per minute real time. Arm and leg activity counts were highly correlated ($r = .69$, $p < .001$), therefore, composite actigraph scores reflecting overall motor activity were calculated by averaging the four limb actigraph scores.

Home AL refers to all activity outside the laboratory. Parents were asked if their children's behavior was typical during the two days in which the actigraphs were worn. In those cases where the parent reported that a child's behavior was not typical (*n*=4), actigraph data was set to missing. In the lab, actigraph scores correlated across the two situations ($r=.53$, $p<.001$), and prior analyses found that the same genetic effects influenced both situations (Saudino & Zapfe, in press) therefore, the two lab measures were combined to create an overall laboratory measure of actigraph AL.

Parent reports of AL—The *Toddler Behavior Assessment Questionnaire* (TBAQ, Goldsmith, 1996) provided a parent-rating measure of AL in the home. The activity subscale of the TBAQ consists of 10 questions regarding the child's behavior in specific situations (e.g., “When playing inside the house, how often did your child climb over furniture?”). Parents indicated on a 7-point scale how frequently the child demonstrated the behavior during the previous month (1=*Never* to 7=*Always*). Internal consistency for this scale was .78.

Observer reports of AL—The *Infant Behavior Record* (IBR, Bayley, 1969) provided an observational measure of AL in the laboratory. The IBR consists of 30 items, 25 of which are 5- or 9-point rating scales evaluating broad dimensions of infant behavior, including interpersonal, affective, motivational, and sensory behavioral domains. Factor analyses of the IBR has yielded three temperament factors: activity, task orientation, and affect-extraversion (Matheny, 1980). The activity factor includes the infant's general level of body motion and degree of energy exhibited during the assessment. Testers completed the IBR following both lab visits. The standardized unweighted items were aggregated and averaged across visits. This composite score displayed good reliability in terms of internal consistency (α=.90), inter-rater agreement (r=.71, p< .001).

Data Transformation

Consistent with prior research, actigraph scores were positively skewed (Saudino & Eaton, 1991, 1995; Wood, Saudino, Rogers, Asherson, & Kuntsi, 2007) and were square-root transformed to create a more normal distribution. Because twin covariances can be inflated by variance due to sex, all scores used in the behavioral genetic analyses were residualized for sex effects (see McGue & Bouchard, 1984).

Correlational Analyses

The twin method involves comparing genetically identical (MZ) twins with fraternal (DZ) twins who share approximately 50% of their segregating genes. Genetic influences are implied when cotwin similarity covaries with the degree of genetic relatedness. If heredity affects a trait, the two-fold greater genetic similarity of MZ twins is expected to make them more similar than DZ twins. Intraclass correlations typically serve as indices of cotwin similarity. An MZ correlation that is greater than the DZ correlation suggests genetic influence on the phenotype. DZ correlations that exceed one-half the MZ correlation suggest the presence of shared environmental influences. Differences within pairs of MZ twins are due to nonshared environmental influences and measurement error.

To evaluate genetic and environmental sources of covariance across methods of assessing AL, cross-twin cross-method correlations were calculated. The cross-twin cross-method correlation involved correlating Twin 1's score for AL using one method with Twin 2's score for AL using another method and vice versa. Genetic contributions to the covariance between two measures are implied when the MZ cross-twin correlation is greater than the DZ cross-twin correlation.

Model-Fitting Analyses

Although correlational analyses can give an impression of genetic and environmental sources of variance and covariance, model-fitting analyses are required to estimate the size of these effects. Bivariate correlated factors models were used to estimate genetic and environmental variances for each measure and to explore the extent to which different measures of AL within the same situation tap the same genetic and environmental effects. Models were fit to raw data using a maximum likelihood pedigree approach implemented in Mx structural equation modeling software (Neale, Boker, Xie, & Maes, 2006). This approach allows the inclusion of participants with incomplete data. The overall fit of a model can be assessed by calculating twice the difference between the negative log-likelihood (-2LL) of the model and that of a

saturated model (i.e., a model in which the variance/covariance structure is not estimated and all variances and covariances for MZ and DZ twins are estimated). The difference in -2LL is asymptotically distributed as χ^2 with degrees of freedom equal to the difference in the number of parameters in the full model and that in the saturated model.

The model, depicted as a path diagram in Figure 1, partitions the phenotypic covariance between the actigraph home measure and the TBAQ into genetic, shared and nonshared components. The latent variables A1, C1, and E1 refer to the genetic (additive), shared and nonshared environmental influences on one AL measure (e.g., actigraph) and A2, C2, and E2 refer to the genetic and environmental influences on a second measure of AL (e.g., TBAQ). The path coefficients, *h*, *c*, and *e*, are standardized partial regressions indicating the relative influence of the latent variables on the phenotypes. The square of these path coefficients estimates the genetic and environmental variances for that measure. Of particular interest in this model are the estimated parameters r_g, r_c and r_e; the genetic, shared environmental, and nonshared environmental correlations, respectively, between AL measures. The genetic correlation indicates the extent to which genetic effects on one measure correlate with genetic effects on another, *independent of the heritability of each measure*. The genetic factors that influence two measures can covary perfectly even though the genetic factors on each measure contribute only slightly to the phenotypic variance. Thus, r_g can be 1.0 even though the genetic contribution to the phenotypic correlation is only modest if the heritability of each measure is modest and the same genetic effects operate on each measure. Conversely, two measures may be substantially heritable, but the genetic correlation would be zero if the genetic effects on the two measures do not overlap. Similar logic applies to r_c and r_e. An estimate of *bivariate heritability*, the proportion of the phenotypic correlation that is due to genetic factors, can also be obtained from this model.

Results

Descriptive Statistics

Table 1 lists the means and standard deviations by sex and zygosity. Mean differences were evaluated using generalized estimating equations (GEE) implemented in the SAS GENMOD procedure to account for dependence in the data due to the fact that our sample comprised twin pairs. GEE are an extension of the standard generalized linear models that allow modeling of correlated data (Liang & Zeger, 1986;Zeger & Liang, 1986). Consistent with previous research (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006), males were significantly more active than females (Actigraph Home: z=-2.32, p<.05; TBAQ: z=-4.17, p<.001; Actigraph Lab: z=-4.92, p<.001; IBR z=-5.77, p<.001). Effect sizes ranged from .24 to .53, indicating that the means for males and females differed by approximately one-quarter to one-half of a standard deviation. More important to our twin analyses, MZ and DZ twins did not significantly differ in AL (Actigraph Home: z=-0.74, p=.46; TBAQ: z=-0.31, p=.75; Actigraph Lab: z=0.01, p=.99; IBR z=0.05, p=.96).

Phenotypic Correlations

Parent ratings of AL on the TBAQ were modestly correlated with actigraph AL in the home (r=.25, p<.01). However, the phenotypic correlation between actigraph AL and observer ratings of AL in the lab was substantial (r=.67, p<.001).

Twin Correlations

For all measures of AL, intraclass correlations for MZ twins exceeded those for DZ twins, suggesting genetic influences on AL (Table 2). The very high intraclass correlations for actigraph AL in the home for both MZ and DZ twins suggest that shared environmental influences also contribute to co-twin resemblance for this measure. Cross-twin cross-method

correlations for MZ twins were consistently higher than those for DZ twins, indicating that genetic factors contribute to the phenotypic correlation between different measures of AL.

Model-fitting Analyses

Table 3 presents the fit statistics for the model-fitting analyses. In addition to the full model, a series of reduced models were fit to the data to evaluate sources of covariance between measures. For AL measures in the home, the full ACE model fit the data well, however, it was possible to drop both the shared and nonshared environmental correlations between AL measures (i.e., r_c and r_e) from the model without a significant decrement in fit. In contrast, dropping the genetic correlation (r_g) resulted in a significantly worse fit. Thus, only genetic factors significantly contributed to the covariance between the actigraph and parent ratings of AL in the home.

For AL measures in the lab, the full ACE model also provided a good fit but the best fitting model included only genetic and nonshared environmental contributions to the covariance between the actigraph and observer ratings on the IBR. The shared environmental correlation could be dropped from the model without worsening the fit to the data.

Estimates of genetic and environmental variances and correlations from the best-fitting models are presented in Table 4. All measures were significantly heritable, however, the magnitude of genetic effects varied across measures. For example, in the home, genetic influences accounted for approximately one-third of the variance in actigraph AL, and over three-quarters of the variance in parent-rated AL on the TBAQ. The nonoverlapping confidence intervals for these two heritiabilities suggest that this difference is significant. Actigraph AL in the home was the only measure to display significant shared environmental influences, consequently it was not surprising to find that shared environmental influences do not contribute to the covariance between measures—both measures must have shared environmental variance if shared environments contribute to the covariance between measures. The genetic correlation of .38 between the actigraph in the home and parent-rated AL on the TBAQ suggests moderate genetic overlap across the two measures. Nonetheless, this moderate genetic covariance explains all of the phenotypic covariance between the actigraph and the TBAQ (i.e., bivariate heritability was 1.0).

AL measures in the lab show a different pattern of results. The heritiabilities of both the actigraph and the observer-ratings on the IBR were moderate and confidence intervals were overlapping. Moreover, the genetic correlation of .95 indicates that both lab-based measures of AL are essentially influenced by the same genetic effects. There was also moderate overlap in nonshared environmental influences on both measures (r_e=.43). The bivariate heritability for the actigraph and IBR was .72 (CI: .60-.82), indicating that genetic factors explain 72% of the phenotypic correlation between the two measures. The remaining proportion of covariance (28%, CI: .18-.40) was explained by nonshared environmental factors.[1]

[1]Estimates of r_g based on analyses of cross-correlations do not completely match those from the model-fitting analyses. There are two reasons for this. First, our sample size is relatively small owing to our detailed assessment of AL in the lab and the home. As a result, correlations have large confidence intervals. Second, analyses of cross-correlations do not simultaneously take into account the heritabilities of and the phenotypic covariances between each measure, nor do they consider MZ and DZ differences in variances and covariances (Eaves, Eysenck, & Martin, 1989). Although less powerful, correlational analyses lead to similar conclusions about genetic overlap across measures. There was no significant difference between the MZ and DZ actigraph-parent rating cross correlations, indicating no significant genetic covariance between these measures. In contrast, the difference between the MZ and DZ cross correlations for actigraph and observer ratings was significant and indicates significant genetic covariance.

Discussion

Consistent with prior research, AL as assessed via parents, observers, and actigraphs was genetically influenced. At question, however, was whether different methods of assessing AL tapped the same genetic influences. To some extent, there is genetic overlap between actigraph and rating-based measures of AL, but the degree of overlap differed for parent versus observer ratings.

The modest phenotypic correlation between parent-rated AL and actigraph AL in the home suggests that different factors influence the two AL measures. This was indeed the case. There was no covariance between environmental effects influencing parent-rated and actigraph AL. Genetic factors explained all of the phenotypic correlation, but the genetic correlation between parent-rated and actigraph AL was only moderate. Approximately 40% of the genetic effects on parent-rated AL overlap with those on actigraph AL. Thus, there is a considerable amount of genetic variance that is unique to each measure (i.e., 60%). The finding of very high bivariate heritability might seem puzzling in light of the moderate genetic correlation between the two AL measures, however, this will be the case whenever two variables are heritable but only modestly correlated. In other words, despite the fact that genetic factors mediate the phenotypic association between the two variables, there is substantial genetic variance on parent-rated AL that is independent of genetic variance on actigraph AL, and visa versa. Because both measures assessed AL in the same situation, the independent genetic variances reflect method-specific, not situation-specific, effects. Thus, to some extent, the two methods of assessing AL do not tap the same genetic effects. These results are remarkably similar to Woods et al. (2008) who found a genetic correlation of .39 and bivariate heritability of .95 between parent-rated hyperactivity and actigraph AL in the lab. The current study adds resolution to Woods et al. by ruling out situational differences as an explanation for independent genetic influence on the two measures of AL. Moreover, the similarity in findings across different age groups (early childhood and middle childhood) attests to the robustness of the effects.

A possible limitation to this finding should be acknowledged. Although both parent ratings and actigraphs were used to assess AL in the home, the time frames for the two measures differed. The TBAQ asks parents about activity behaviors observed within the month, whereas the actigraph data was based on a 2-day data collection period. However, the analyses only included actigraph data for those children whose parents indicated that child behavior during data collection was typical of their usual behavior. Moreover, a previous study in which parents' ratings on the TBAQ and mechanically-assessed AL were obtained for *same 48-hour period* (Saudino & Eaton, 1995) produced a similarly low phenotypic correlation between AL measures (r=.14). Analyses of genetic covariance were not conducted on this early data, but analyses based on a small sample of older twins that also included parent ratings and mechanical measures of AL for the same 48-hour period yielded results that were remarkably consistent to the present findings—both in terms of genetic contributions to the phenotypic correlation and substantial unique genetic variance (Saudino, Thompson, & Gagne, 2000). Thus, it is unlikely that the present findings of low phenotypic covariance and genetic independence between parent-rated and actigraph AL are simply due to differences in time frames for behavioral sampling.

In contrast to parent ratings, observer ratings of AL were highly correlated with actigraph AL in the lab. In fact, the phenotypic correlation between the two methods was almost as high as the estimate of inter-rater agreement for the observed measure. Although the phenotypic correlation was largely due to genetic effects, nonshared environmental factors also contributed to the covariance between the two methods. The very high genetic correlation between methods indicates that genetic influences on observer ratings and actigraph measures of AL are almost entirely overlapping. That is, the two measures generally tap the same genetic effects. There

was also moderate overlap in nonshared environmental variances between observer-rated and actigraph AL. Nonshared environmental variance comprises environmental factors that are unique to each member of a twin pair and that make twins different from each other. Twins were seen by different testers which could explain the finding of nonshared environmental covariance across the two methods. Testers followed standardized protocols, however, tester personality and behavioral standards may have influenced their interactions and consequently the AL of the twin tested.

The finding of substantial genetic independence between actigraph AL and parent ratings, but not for actigraph AL and observer ratings, implies that parent and observer ratings of AL are also largely genetically distinct, but this remains an empirical question. The difference in genetic associations between actigraphs and parent and observer ratings may reflect conceptual differences across the measures. Parent ratings on the TBAQ include information about the quality of the child's movement (i.e., running, climbing, squirming) whereas the observer ratings on the IBR focus on the frequency of movement in the lab. This latter view of AL is more in line with the actigraph. Parent ratings of AL have also been shown to be prone contrast effects—rater biases that magnify existing behavioral differences between co-twins and thereby inflate heritability estimates (Saudino et al., 2004). The pattern of twin correlations for the TBAQ in the present study is not consistent with contrast effects, nonetheless, researchers need to carefully consider contrast biases whenever parent-rating measures are employed to assess temperament.

These results have important implications for developmental researchers. Despite the common assumption that parent ratings and more objective measures of temperament tap the same constructs, the present findings indicate that these methods can differ at the level of etiology and are not interchangeable. Findings with one method may not generalize to another, not because of contextual factors, but because different methods engage different processes. This is particularly important for molecular genetic research that seeks to find genes associated with temperament. For example, the fact that to some extent, different methods tap different genetic influences on AL means that researchers will need to consider the method used to assess AL. Failures to replicate molecular genetic findings may be due to method differences between studies.

Acknowledgments

The BUTP is supported by grant MH062375 from the National Institute of Mental Health. The author gratefully acknowledge the parents and twins in the BUTP.

References

Actical Physical Monitoring System Instruction Manual. Bend, OR: Mini Mitter Company, Inc.; 2005.

Bayley, N. Manual for the Bayley Scales of Infant Development. New York: The Psychological Corporation; 1969.

Bayley, N. Bayley Scales of Infant Development. Vol. 2. San Antonio, TX: The Psychological Corporation; 1993.

Campbell SB, Breaux AM. Maternal ratings of activity level and symptomatic behaviours in a nonclinical sample of young children. Journal of Pediatric Psychology 1983;8:73–82. [PubMed: 6842352]

Eaves, LJ.; Eysenck, HJ.; Martin, NG. Genes, Culture and Personality: An Empirical Approach. New York: Academic Press; 1989.

Else-Quest NM, Hyde JS, Goldsmith HH, Van Hulle CA. Gender Differences in Temperament: A Meta-Analysis. Psychological Bulletin 2006;132(1):33–72. [PubMed: 16435957]

Fagan J. The interaction between child sex and temperament in predicting behavior problems of preschool-age children in day care. Early Child Dev Care 1990;59:1–9.

Fagot BI, O'Brien M. Activity level in young children: Crossage stability, situational influences, correlates with temperament, and the perception of problem behaviors. Merrill Palmer Quarterly 1994;40:378–398.

Goldsmith HH. Studying temperament via construction of the Toddler Behavior Assessment Questionnaire. Child Development 1996;67:218–235. [PubMed: 8605830]

Goldsmith HH, Lemery KS, Buss KA, Campos JJ. Genetic analyses of focal aspects of infant temperament. Developmental Psychology 1999;35:972–985. [PubMed: 10442866]

Goldsmith, HH.; Reilly, J.; Lemery, KS.; Longley, S.; Prescott, A. The laboratory temperament assessment battery-preschool version: Description of procedures. University of Wisconsin; 1995.

Halverson CF, Waldrop MF. Relations between preschool activity and aspects of intellectual and social behavior at age 7½. Developmental Psychology 1976;12:107–112.

Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13–22.

Martin, R.; Olejnik, S.; Gaddis, L. Is temperament an important contributor to schooling outcomes in elementary school? Modeling effects of temperament and scholastic ability on academic achievement. In: Carey, W.; McDevitt, S., editors. Prevention and Early Intervention: Individual Differences as resk factors for the mental health of children. New York: Brunner/Mazel; 1994. p. 59-68.

Matheny AP Jr. Bayley's infant behavior record: Behavioral components and twin analyses. Child Development 1980;51:1157–1167. [PubMed: 7193557]

McGue M, Bouchard TJ. Adjustment of twin data for the effects of age and sex. Behavior Genetics 1984;14:325. [PubMed: 6542356]

Mehregany DV. The relation of temperament and behavior disorders in a preschool clinical sample. Child Psychiatry Hum Dev 1991;22:129–136. [PubMed: 1800024]

Neale, MC.; Boker, SM.; Xie, G.; Maes, HH. Mx: Statistical modeling. Vol. 7th. Richmond, VA: Virgina Commonwealth University; 2006.

Pate RR, Almeida MJ, McIver KL, Pfeiffer KA, Dowda M. Validation and Calibration of an Accelerometer in Preschool Children. Obesity 2006;14:2000–2006. [PubMed: 17135617]

Philips K, Matheny AP. Evidence for genetic influence on both cross-situation and situation-specific components of behavior. Journal of Personality and Social Psychology 1997;73:129–138. [PubMed: 9216081]

Price TS, Freeman B, Craig I, Petrill SA, Ebersole L, Plomin R. Infant zygosity can be assigned by parental report questionnaire data. Twin Research 2000;3:129. [PubMed: 11035484]

Rothbart, MK.; Derryberry, D.; Hershey, K. Stability of temperament in childhood: Laboratory infant assessment to parent report at 7 years. In: Molfese, V.; Molfese, D., editors. Temperament and personality development across the lifespan. Hillsdale, NJ: Erlbaum; 2000. p. 85-120.

Rothbart, MK.; Derryberry, D.; Posner, M. A psychobiological approach to the development of temperamen. In: Bates, JE.; Wachs, TD., editors. Temperament: Individual differences at the interface of biology and behavior. Washington, DC: APA Books; 1994. p. 83-116.

Sanson A, Smart D, Prior M, Oberklaid F. Precursors of hyperactivity and aggression. Journal of the American Academy of Child and Adolescent Psychiatry 1993;32:1207–1216. [PubMed: 8282666]

Saudino KJ, Eaton WO. Infant temperament and genetics: An objective twin study of motor activity level. Child Development 1991;62:1167. [PubMed: 1756660]

Saudino KJ, Eaton WO. Continuity and change in objectively assessed temperament: A longitudinal twin study of activity level. British Journal of Developmental Psychology 1995;13:81.

Saudino KJ, Plomin R, DeFries JC. Tester-related temperament at 14, 20, and 24 months: Environmental change and genetic continuity. British Journal of Developmental Psychology 1996;14:129–144.

Saudino KJ, Thompson LA, Gagne JR. Genetic influences on activity level in adolescence: Do the same genetic effects operate on parent ratings, self-ratings and mechanical measures? Behavior Genetics 2000;30:405–406.

Saudino KJ, Wertz AE, Gagne JR, Chawla S. Night and Day: Are Siblings as Different in Temperament as Parents Say They Are? Journal of Personality and Social Psychology 2004;87:698–706. [PubMed: 15535780]

Saudino KJ, Zapfe JA. Genetic influences on activity level in early childhood: Do situations matter? Child Development. in press

Seifer R, Sameroff AJ, Barrett LC, Krafchuck E. Infant temperament measured by multiple observations and mother reports. Child Development 1994;65:1487–1490.

Teglasi H, MacMahon BH. Temperament and common problem behaviors of children. Journal of Applied Developmental Psychology 1990;11:331–349.

Wood AC, Rijsdijk F, Saudino KJ, Asherson P, Kuntsi J. High heritability for a composite index of children's activity level measures. Behavior Genetics 2008;38(3):266–276. [PubMed: 18297388]

Wood AC, Saudino KJ, Rogers H, Asherson P, Kuntsi J. Genetic influences on mechanically-assessed activity level in children. Journal of Child Psychology and Psychiatry 2007;48:695–702. [PubMed: 17593150]

Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 1986;42:121–130. [PubMed: 3719049]

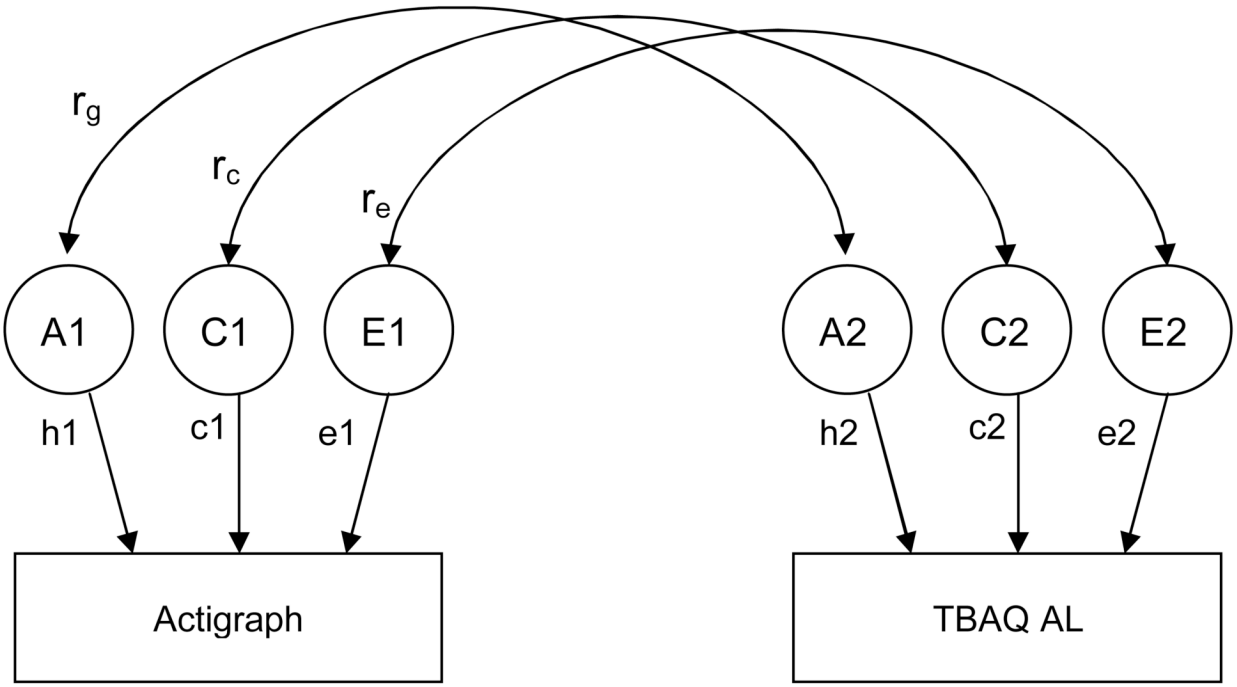

Figure 1.
Correlated factors model.

Table 1

Means (and Standard Deviations) for Measures of AL by Sex and Zygosity

	Males		Females		Effect Size[a]	
AL Measure	**MZ**	**DZ**	**MZ**	**DZ**	**Sex**	**Zygosity**
Home						
Actigraph	21.63 (2.49)	21.21 (2.45)	20.79 (2.45)	20.85 (2.10)	.24	.07
TBAQ	44.42 (9.01)	42.56 (9.41)	39.04 (6.98)	40.54 (8.01)	.41	.01
Lab						
Actigraph	26.97 (5.53)	27.17 (5.02)	24.74 (5.02)	24.54 (4.83)	.47	.02
IBR	0.18 (0.81)	0.23 (0.81)	-0.20 (0.77)	-0.24 (0.71)	.53	.01

Note. MZ=monozygotic twins, DZ=dzygotic twins.

[a]Cohen's *d*.

Table 2

Twin Intraclass Correlations, Cross-Method Correlations (and 95% Confidence Intervals).

	Intraclass Correlations		Cross Correlations	
AL Measure	**MZ**	**DZ**	**MZ**	**DZ**
Home				
Actigraph	.87 (.83-.91)	.70 (.61-.77)		
TBAQ	.81 (.75-.86)	.46 (.33-.57)	.29 (.13-.43)	.12 (-.03-.27)
Lab				
Actigraph	.63 (.52-.72)	.36 (.22-.49)		
IBR	.54 (.41-.65)	.38 (.24-.50)	.47 (.34-.59)	.28 (.13-.41)

Note. MZ=monozygotic twins, DZ=dzygotic twins.

Table 3

Model-fitting Statistics for Bivariate ACE Model

Variable and Model	Overall Fit of Model[a] -2LL	*df*	χ^2	*df*	*p*	AIC	Relative Fit of Model[b] $\Delta\chi^2$	*df*	*p*
Home									
Actigraph - TBAQ									
Saturated	6669.795	1208							
Full	6676.629	1219	6.384	11	.812	-15.166			
Drop r_g	6680.494	1220	10.699	12	.555	-13.301	3.865	1	.049
Drop r_c	6678.523	1220	8.728	12	.726	-15.272	1.894	1	.325
Drop r_e	6676.632	1220	6.837	12	.867	-17.163	0.002	1	.964
Drop r_c & r_e	**6678.573**	**1221**	**8.778**	**13**	**.789**	**-17.221**	**1.944**	**2**	**.378**
Lab									
Actigraph - IBR									
Saturated	4777.183	1220							
Full	4790.314	1231	13.131	11	.285	-8.869			
Drop r_g	4797.262	1232	20.079	12	.070	-3.921	6.948	1	.008
Drop r_c	**4791.416**	**1232**	**14.233**	**12**	**.289**	**-9.767**	**1.102**	**1**	**.294**
Drop r_e	4835.375	1232	58.192	12	.000	34.192	45.061	1	.000

Note. -2LL=Likelihood Statistic. χ^2=Chi-square fit statistic. *df*=Degrees of freedom. AIC=Akaike's Information Criterion. r_g=genetic correlation, r_c=shared environmental correlation, r_e=nonshared environmental correlation. Best fitting model indicated in bold.

[a]Overall fit of the model is determined by the difference in -2LL of the model and that of a saturated model.

[b]Relative fit of the model determined by the χ^2 difference ($\Delta\chi^2$) between full bivariate ACE model and reduced model.

Table 4

Estimates of Genetic and Environmental Variances and Genetic and Environmental Correlations (and 95% Confidence Intervals) from the Best-Fitting Model

	Variance Components			Covariances		
AL Measure	a^2	c^2	e^2	r_g	r_c	r_e
Home						
Actigraph	.32 (.18-.48)	.54 (.39-.66)	.14 (.11-.18)			
TBAQ	.82 (.61-.87)	.01 (.00-.22)	.17 (.13-.22)	.38 (.22-.56)	.00	.00
Lab						
Actigraph	.59 (.37-.70)	.03 (.00-.21)	.38 (.30-.48)			
IBR	.40 (.25-.62)	.13 (.00-.25)	.46 (.37-.57)	.95 (.72-1.0)	.00	.43 (.30-.54)

Note. a^2=genetic variance, c^2=shared environmental variance, e^2=nonshared environmental variance. r_g=genetic correlation, r_c=shared environmental correlation, r_e=nonshared environmental correlation.