Published in final edited form as:

Child Dev. 2012 January ; 83(1): 266–281. doi:10.1111/j.1467-8624.2011.01680.x.

Sources of Continuity and Change in Activity Level in Early Childhood

Kimberly J. Saudino
Boston University

Abstract

Actigraphs and parent and observer ratings were used to explore genetic influences on continuity and change in activity level (AL) in early childhood. Over 300 pairs of twins wore actigraphs for a 48-hour period in the home and laboratory at ages 2 and 3. AL was genetically influenced at both ages with little evidence of differential heritability across age. For all measures, genetic influences contributed to phenotypic continuity. With the exception of the actigraph measure of AL in the home, new genetic effects emerged at age 3 indicating that genetic factors influence both continuity and change in AL in early childhood. Nonshared environmental influences were also a source of change in AL across the transition from infancy to early childhood.

Keywords

twins; activity level; genetic; development; model fitting

Activity level (AL) is a focal behavioral dimension that appears in most major theories of infant and child temperament (e.g., Goldsmith et al., 1987; Goldsmith, Lemery, Buss, & Campos, 1999). Defined as an individual's customary level of energy expenditure through gross motor movement, AL in infancy and early childhood has been linked to a wide variety of developmental outcomes including externalizing behaviors (Fagan, 1990; Fagot & O'Brien, 1994; Gjone & Stevenson, 1997; Teglasi & MacMahon, 1990); hyperactivity (Campbell & Breaux, 1983; O'Donnell & Van Tuinan, 1979; Schaugency & Fagot, 1993); poor adjustment to daycare or school (Jewswan, Luster, & Kostelnik, 1993; Klein & Ballantine, 1988; Zajdeman & Minnes, 1991); and conflict and negativity within sibling-, peer- and parent-child interactions (Brody, Stoneman, & Burke, 1987; Guralnick & Groom, 1990; Katainen, Räikkönen, & Keltikangas-Järvinen, 1997, 1998; Laible, Panfile, & Makariev, 2008; Olson, Bates, Sandy, & Lanthier, 2000).

Developmental research finds that AL changes across age. An analysis of motion recorder data from 12 studies across the lifespan found that AL increases linearly from infancy to middle childhood (Eaton, McKeen, & Campbell, 2001). Although age-related change in mean levels of AL are interesting, stability or rank-order change across age may be more important given AL's significance as an individual differences variable. From an individual differences perspective, developmental change in AL is substantial. A meta-analysis of longitudinal studies of personality and temperament found that most temperament dimensions displayed moderate stability across infancy and childhood (ranging from .35 to .47), however, the stability of AL was more modest (.28) (Roberts & DelVecchio, 2000). Thus, in terms of individual differences, although there is some continuity in AL in

childhood, change predominates. At question then, is what factors explain individual-level continuity and change in activity?

One possible explanation of individual developmental change in AL is genes. Genes are dynamic in nature, changing in the quantity and quality of their effects as the organism changes developmentally (Plomin & Nesselroade, 1990). Moreover, different genes can be involved at different stages of development. Thus, genetic factors can be a source of change as well as continuity in behavioral development. Temperament theories posit that temperament dimensions have a biological or constitutional basis (Goldsmith et al., 1987). Twin studies employing parent ratings, observer ratings and mechanical measures of AL have consistently found that individual differences in AL in infancy, early childhood and middle childhood are, in part, genetically influenced (e.g., Braungart, Plomin, DeFries, & Fulker, 1992; Cyphers, Phillips, Fulker, & Mrazek, 1990; Saudino, 2009; Saudino & Eaton, 1991; Stevenson & Fielding, 1985; Wood, Saudino, Rogers, Asherson, & Kuntsi, 2007); however, the fact that AL is heritable at different ages does not mean that the same genetic influences operate at each age. That is, the genes that influence AL at one age may differ from those that influence AL at another. This is true even when estimates of heritability are equivalent across age. Similarly, differences in the magnitude of heritability estimates across age does not mean that the genes that operate on AL differ from one age to the next—only that there is a difference in the extent to which genetic factors contribute to behavioral variability in AL. Thus, comparing heritability estimates across age cannot inform about sources of continuity and change and do not address developmental processes.

From a developmental perspective, the measurement of within-person change is more interesting than cross-sectional age differences because the measurement of change is more informative about underlying processes (Wohlwill, 1973). For the same reason, genetic influences on individual continuity and change are more interesting than the presence or absence of genetic influences at various ages. Studies examining differential heritability across age are essentially cross-sectional and do not address *mechanisms* of change. The question of genetic and environmental influences on individual developmental change can be addressed by assessing genetic contributions to phenotypic continuity and change across age (i.e., age-to-age stability and instability). Analyses of genetic contributions to phenotypic continuity permit the estimation of the extent to which genetic effects on a trait at one age overlap with genetic effects at another age (i.e., the genetic correlation) and, further, whether new genetic influences on the trait emerge across time. Such analyses also inform about environmental sources of individual continuity and change and can therefore provide important information about developmental *processes*.

Longitudinal studies with genetically-informative samples (i.e., twin or adoption designs) are needed to inform about genetic and environmental contributions to individual continuity and change. However, despite the fact that AL is a developmentally significant behavioral dimension, few studies have used these designs to examine the etiology of individual developmental change in AL. In an early attempt to explore genetic influences on individual developmental change in temperament, the Louisville Twin Study examined twin concordances for age-to-age change profiles for laboratory observations of activity across 6 to 18 months, and 12 to 24 months of age (Matheny, 1989; Wilson & Matheny, 1986). MZ twins demonstrated greater resemblances than DZ twins on age-to-age change profiles suggesting that across infancy and early childhood the developmental patterns for AL are, in part, regulated by genetic influences. A problem with the analysis of change profiles, however, is that they do not reveal the extent to which the genetic effects that influence temperament are consistent from one age to another (Plomin & Nesselroade, 1990).

To date, the MacArthur Longitudinal Twin Study (MALTS; Emde & Hewitt, 2001; Plomin et al., 1990) is the only large-scale study to directly examine genetic and environmental contributions to phenotypic continuity in AL (Saudino & Cherny, 2001a, 2001b; Saudino, Plomin, & DeFries, 1996). The results differ according to method of assessing AL. Observer ratings of AL in the lab at 14, 20, 24 and 36 months of age yielded modest age-to-age stability (i.e., correlations ranging from .14 to .27) indicating considerable individual developmental change (Saudino & Cherny, 2001b). Age-to-age *genetic* correlations from 14 to 24 months approached unity, suggesting that there is substantial overlap between the genetic effects on AL across the second year of life. It was these overlapping genetic effects that accounted for the modest phenotypic stability in AL. Moreover, from 14 to 24 months of age, genetic factors contributed only to continuity in AL whereas developmental change was due entirely to nonshared environmental influences. However, at 36 months of age, new genetic effects on AL emerged and age-to-age genetic correlations between 36 months and earlier ages were only moderate (i.e., .31-.55). The finding of novel genetic effects at age 3 suggests that genetic factors contribute to individual developmental change in AL across the transition from infancy to early childhood. These genetic influences on individual developmental change are intriguing, but they are limited by methodological issues. That is, genetic change in AL at age 3 coincided with changes in content of the measure used to assess AL raising the question as to whether the observed genetic change reflects true developmental change or differences in the measure at 36 months.

Parent ratings of AL on the Colorado Child Temperament Inventory (CCTI) in MALTS produced a different pattern of results. Consistent with previous research (Plomin & DeFries, 1985; Rothbart, 1986; Wilson & Matheny, 1986), parent ratings of AL displayed much higher phenotypic stabilities (i.e., in the range of .58 to .70) than observational measures. Moreover, unlike the stability of observer rated-AL, which was entirely due to genetic factors, the stability of parent ratings was mediated by genetic *and* nonshared environmental influences (Saudino & Cherny, 2001a). Nonshared environmental influences are those environmental factors that are unique to each individual that make family members different from each other. The finding that nonshared environmental factors contribute systematically to stability for parent ratings but not for observational measures of AL suggests the presence of stable parent rating biases that exaggerate behavioral differences between siblings (Plomin et al., 1993). These biases are referred to as “contrast effects” and have been shown to influence the CCTI ratings in the MALTS dataset (Saudino, Cherny, & Plomin, 2000). Thus, the MALTS findings suggest that the stability of parent ratings reflect not only stable genetic influences on AL, but also stable parental perceptions about differences in their twins' behavior (i.e., the tendency for parents to label one twin as active and the other as inactive might persist across age intervals from 14 to 36 months).

Overall, the MALTS data yield inconclusive results regarding sources of individual continuity and change in AL. When assessed via behavioral observation, developmental change is confounded with changes in the properties of the measures used to assess AL; and when assessed via parent ratings, continuity in AL may be confounded with continuity in rater contrast biases. New measurement strategies are needed to better understand the factors that influence developmental continuity and change in AL. In the present study motion recorders (actigraphs) were used to record the AL of twins at 2 and 3 years of age—the ages at which the MALTS observational data suggests that new genetic effects on AL may emerge. Actigraphs have been shown to be valid and reliable instruments for recording the activity of young children (Freedson et al., 1997; Pate, Almeida, McIver, Pfeiffer, & Dowda, 2006). Such measures are more objective, permitting the transition away from human inferences based on qualitative behavioral definitions to a more quantitative measure of the physical forces that are associated with human activity, thus addressing the problems of differing item content or rater biases. This first large-scale longitudinal twin study of

mechanically-assessed AL in early childhood provides a more rigorous test of the genetic and environmental factors that influence individual developmental change in AL during early childhood. Based on previous research, it was predicted that AL will be genetically influenced at both ages and that genetic factors will contribute to both individual continuity and change across age.

Because prior research has indicated that genetic effects on AL may vary across situations (Saudino & Zapfe, 2008), the present study examines individual developmental change in the home and lab separately. The issue of situational differences in the etiology of development in AL (or any other temperament dimensions) has not been previously examined, but it is possible that the sources of developmental continuity and change in AL may differ in the two situations. Similarly, because different methods of assessing AL may be influenced by different genetic effects (Saudino, 2009), observer and parent ratings of AL were also included in the study. In contrast to MALTS, the observer rating measures in the present study did not differ across age and the parent-rating measure employed, the *Toddler Behavior Assessment Questionnaire* (Goldsmith, 1996), has been shown to be less prone to rater contrast effects (Goldsmith, Buss, & Lemery, 1997). The inclusion of these more traditional temperament-based measures of AL in the study allows us to examine whether the findings generalize across measures.

Method

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. Three hundred and fourteen same-sex pairs of twins (144 MZ, 168 DZ) participated in the age 2 assessments, and 304 of these twin pairs (141 MZ, 163 DZ) returned for the age 3 assessments (96.8% retention rate). 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 and third birthdays. At each age, the procedure consisted of two visits, 48-hours apart, to the laboratory. Actigraphs were attached to each child (see below) at the first visit and removed at the second visit (i.e., the twins wore the actigraphs home and continuously during the 2-day interval between lab visits). At the initial visit, informed consent from parents was obtained. After attachment of the actigraphs, one twin was assessed within a standardized test situation, while the other twin was assessed within a laboratory play situation. The order of situations was counterbalanced across first- and second-born twins (i.e., for 50% of the sample the first-born twin was in the play situation on visit 1 and the test situation on visit 2; and for 50% of the sample the second-born twin was in the play situation on visit 1 and the test situation on visit 2). The test situation involved administration of the Mental Scale of *Bayley Scales of Infant Development—Second Edition* (Bayley, 1993). The play situation comprised activity

and inhibitory control episodes (arc of toys, corral of balls, workbench, fidget video, dinky toys, snack delay, gift) 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, within age for each twin the tester was the same across the two laboratory situations.

Measures

Mechanical assessment of AL—AL was continuously assessed for 48 hours with the Minimitter actical (actigraph). This device is a miniature omnidirectional accelerometer that has been designed to detect low frequency (0.5–3.2 Hz) G-forces (0.05–2.0 Hz) within the range of normal human movement (Heil, 2006). Activity is sampled 32 times per second (32Hz) and is expressed as activity counts representing the frequency and amplitude of acceleration events occurring over a 1-minute measurement epoch (Actical Physical Monitoring System Instruction Manual, 2005). Thus, activity counts as assessed with the Actical actigraph correspond to changes in physical activity energy expenditure (Heil, 2006). Because actigraphs collect real-time data, it is possible to extract activity counts for specific time periods (i.e., home and lab).

Motion recorders have been shown to be valid and reliable instruments for recording the activity of infants (Eaton, McKeen, & Lam, 1988), toddlers (Saudino, 2009), and preschool children (Finn & Specker, 2000; Pate et al., 2006). For example, actigraph counts were significantly correlated with direct observations of activity in toddlers and preschoolers (Finn & Specker, 2000; Saudino, 2009) and with oxygen consumption, a metabolic measure of energy expenditure, in preschoolers (Pate et al., 2006). In school-aged children, actigraphs have been shown to differentially discriminate between children diagnosed with ADHD and other diagnostic groups or controls (e.g., Hall, Halperin, Schwartz, & Newcorn, 1997; Matier-Sharma, Perachio, Newcorn, Sharma, & Halperin, 1995; McGrath, Handwerk, Armstrong, Lucas, & Friman, 2004) and are sufficiently sensitive to detect changes in activity following drug treatment for ADHD and across situations (e.g., Konrad, Gunther, Heinzel-Gutenbrunner, & Herpertz-Dahlmann, 2005).

Four randomly-selected actigraphs were assigned to each twin, one for each limb. Assignment of actigraph to limb was also random. Actigraphs were attached by means of tyvek® adhesive or plastic snap 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 a sheet for logging the times when the actigraphs were off a limb. To differentiate actigraphs within twin pair, each 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 engage in normal activities and to maintain daily routines with their children during the 2-day data collection period. 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 (Age 2 $r = .69$, $p < .001$; Age 3 $r = .73$, $p < .001$), therefore, composite actigraph scores reflecting overall motor activity were calculated by averaging the four limb actigraph scores.

Parents were asked if their children's behavior was typical during the two days in which the motion recorders were worn. In those cases where the parent reported that a child's behavior

was not typical ($n=4$), actigraph data were not used. Actigraph scores correlated substantially across the lab and play situations (Age 2 $r=.64$, $p<.001$; Age 3 $r=.58$, $p<.001$), and prior analyses found that the same genetic effects influenced both situations (Saudino & Zapfe, 2008) therefore, the lab measures were combined to create an overall laboratory measure of actigraph AL. The home situation refers to all time outside the laboratory within the 2-day data collection periods.

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 were asked to indicate 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 by Matheny (1980) yielded three factors related to dimensions found in most systems of temperament: Activity, Task Orientation, and Affect-Extraversion. The activity factor includes the child's general level of body motion and degree of energy exhibited during the test situation. Although initially designed for children up to 30 months of age, the IBR has been successfully used to provide an observational measure of temperament in children 3 to 8 years of age (Deater-Deckard, Petrill, & Thompson, 2007; Deater-Deckard, Petrill, Thompson, & DeThorne, 2005, 2006; Saudino & Cherny, 2001b). 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 ($\alpha=.90$), inter-rater agreement ($r=.71$, $p<.001$).

Data Transformation

Consistent with prior research (Saudino & Eaton, 1991, 1995; Wood et al., 2007), actigraph scores in the home and lab at age 2 and the lab at age 3 were positively skewed and were square-root transformed to create a more normal distribution (see Eaton, McKeen, & Saudino, 1996). Because twin covariances can be inflated by variance due to sex, all scores were residualized for sex effects (see McGue & Bouchard, 1984). These residualized scores were used in all behavioral genetic analyses.

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—environmental influences common to both members of a twin pair that enhance cotwin similarity. Differences within pairs of MZ twins are due to nonshared environmental influences and measurement error.

To evaluate genetic and environmental sources of phenotypic continuity across age, cross-twin cross-age correlations were calculated. The cross-twin cross-age correlation involved

correlating Twin 1's score for AL at age 2 with Twin 2's score for AL at age 3 and vice versa. Genetic contributions to the phenotypic continuity 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 across age, model-fitting analyses are required to provide accurate estimates of these 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.

Longitudinal genetic models decompose both the *variance* of a trait at each age and the *covariances* between the trait at multiple ages into additive genetic effects (A), shared environmental effects (C), and nonshared environmental effects (E) (Plomin, 1989). Additive genetic influences refer to the sum of the average effect of all genes that influence a trait. Based on the degree of genetic relatedness, the A factors correlate 1.0 and .5 for MZ and DZ twins, respectively. The C factors refer to the influence of shared rearing environments on twin resemblance. Because all twins were reared in the same family, shared environments correlate 1.0 for both MZ and DZ twins. Finally, the E factors reflect nonshared environmental influences that are unique to each member of a twin pair, including measurement error.

In the present study, a Cholesky decomposition model was used to investigate the sources of continuity and change across age (Neale & Cardon, 1992, see Figure 1). Because prior analyses indicated significant measure- and situation-specific variance (Saudino, 2009; Saudino & Zapfe, 2008), this model was fit separately for each AL measure. The rectangles represent the phenotypic variances of AL at ages 2 and 3. The circles represent the latent genetic, shared environmental, and nonshared environmental variables that may influence AL. A_2 , C_2 , and E_2 are the respective genetic and environmental effects that are common to ages 2 and 3; and A_3 , C_3 , and E_3 represent the genetic and environmental effects that are unique to age 3. The path coefficients are standardized partial regressions indicating the relative influence of the latent variables on the phenotypes. Genetic continuity refers to genetic effects that persist across time. In the model this is suggested when both path coefficients within the common A factor are significant (i.e., a_{21} and a_{22}). Multiplying these two path coefficients provides an estimate of the genetic contributions to the phenotypic across-age correlation (i.e., phenotypic continuity). Estimates of shared and nonshared environmental mediation of phenotypic continuity can be obtained by applying the same principles. Age-to-age genetic correlations (r_G) for a trait can be calculated as the genetic contribution to phenotypic stability divided by the product of the square roots of the heritabilities at each age (Plomin, 1989; Posthuma et al., 2003). r_G indicates the extent to which genetic effects at one age correlate with genetic effects at another *independent of the heritability* of the trait at each age. The genetic factors that influence a trait at two ages can covary perfectly even though the genetic factors at each age contribute only slightly to the phenotypic variance. Thus, r_G can be 1.0 even though the genetic contribution to phenotypic continuity is modest if the heritability at each age is modest and the same genetic effects operate at the two ages. Genetic contributions to change are conceptualized as new genetic effects at one age that are independent of genetic effects at an earlier age. For example, path

a_3 represents new genetic effects that emerge at 3. Similar logic applies to environmental variances and covariances.

For each AL measure, the full ACE model was first fit to the data. A series of reduced models were then fit to see if the models could be simplified. The CE model includes only shared and nonshared environmental effects and the difference between this reduced model and the full ACE model tests the significance of genetic effects on AL. Similarly, the AE model includes genetic and nonshared environmental effects and tests the significance of shared environmental influences. Finally, the E model drops both genetic and shared environmental influences and will only fit the data when there is no significant familial resemblance in AL.

For the parent ratings of AL, an additional model was fit to the data. Because parent ratings of AL have been shown to be prone to rater contrast biases (Saudino et al., 2000; Saudino, McGuire, Reiss, Hetherington, & Plomin, 1995), a sibling interaction model (ACEs) was fit to explicitly test for these effects in the TBAQ data. This model is a modification of the full ACE model to include paths (s in Figure 1) from each twin's phenotype to the phenotype of their cotwin within each age (Neale & Cardon, 1992). Under this model each twin's phenotype is a function of additive genetic effects, shared environmental influences, nonshared environmental influences, and the phenotype of their sibling (i.e., the rating of one twin is influenced by the rating of the other). Sibling interaction effects are indicated if the ACE model fits significantly worse than the ACEs model.

Results

Descriptive Statistics

Table 1 lists the means and standard deviations by age and sex. Mean differences were evaluated using linear mixed models fitted using the SAS PROC MIXED procedure to account for dependence in the data due to the fact that the longitudinal sample comprised pairs of twins nested in families (Singer, 1998). Consistent with previous research, males were significantly more active than females on all measures of AL (Eaton & Enns, 1986; Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006; Saudino & Eaton, 1995). With the exception of observer-rated AL, there were significant age effects for AL. The pattern of change across age, however, differed according to AL measure or situation. For the actigraph measures, AL in the home significantly increased, whereas AL in the lab decreased from age 2 to 3. This latter effect must be interpreted in terms of the significant age by sex interaction. Follow-up tests found that only males showed a significant age effect for actigraph AL in the lab (Males: $F(1, 165)=39.92, p < .001$; Females: $F(1, 145)=0.50, p = .58$). Parent ratings of AL showed a modest decline across age with no age by sex interaction. The age effect for observer-rated AL in the lab approached significance ($p < .09$) and there was a significant age by sex interaction. Males tended to decrease, whereas females increased in observer-rated AL across age, but these effects were not significant in follow-up tests (Males: $F(1, 161)=2.65, p = .11$; Females: $F(1, 144)=2.52, p = .11$).

Phenotypic Continuity and Change

Cross-age stability correlations are presented in Table 2. The correlations are moderate in magnitude indicating both continuity and change in individual differences in AL. Again, consistent with previous research (Plomin & DeFries, 1985; Rothbart, 1986; Wilson & Matheny, 1986), the stability of parent ratings of AL is higher than the more objective measures.

Twin Correlations

For all measures at both ages, the intraclass correlations for MZ twins exceeded those for DZ twins suggesting genetic influences contribute to individual differences in AL in toddlerhood (see Table 3). The very high intraclass correlations for both MZ and DZ twins for the actigraph measures of AL in the home are particularly noteworthy and suggest that shared environmental influences also contribute to cotwin resemblance in this situation. The DZ correlations for observer-rated AL also suggest possible shared environmental influences, but this was not apparent for actigraph-assessed AL in the same situation. Contrast biases do not appear to be influencing parent ratings of AL on the TBAQ. These parent-rating biases yield DZ correlations that are much lower than would be predicted by a genetic hypothesis. That is, a simple genetic hypothesis predicts that the DZ correlation should be approximately one-half the MZ correlation, but when contrast effects are present the DZ correlation is often near zero or negative (Saudino, 2003). This pattern did not emerge for parent-rated AL on the TBAQ. For all AL measures, cross-twin cross-age correlations for MZ twins were consistently higher than those for DZ twins, indicating that genetic factors contribute to phenotypic stability in AL.

Model-fitting Analyses

Table 4 presents the fit statistics for the model-fitting analyses. For the measure of actigraph AL in the home, the full ACE model provided the best fit (i.e., it was not possible to further simplify the model). Both genetic and shared environmental influences contribute significantly to familial resemblance in this measure of AL. For all remaining measures of AL, the AE model provided the best fit to the data. Shared environmental influences could be dropped from the models without a significant decrement in fit. In contrast, dropping genetic effects from the models resulted in significantly worse fits to the data. As suggested by the pattern of correlations for parent-rated AL, there was no indication of contrast effects. The sibling interaction terms were nonsignificant and could be eliminated from the model without worsening the fit to the data.

Estimates of genetic and environmental variances at each age and covariances between ages 2 and 3 are presented in Table 5. The variance components for AL at age 3 are decomposed into variances that persist from age 2 (“Continuity”) and variances that are unique to age 3 (“Change”). For each variance component (i.e., a^2 , c^2 , and e^2) at age 3, summing the values for continuity and change provides an estimate of the total variance for that component. For example, the total genetic variance for actigraph AL in the home at age 3 is $.25 + .00 = .25$, the total shared environmental variance is $.04 + .51 = .55$, and the total nonshared environmental variance is $.00 + .20 = .20$.

There was no evidence of differential heritability across age. Genetic and environmental influences on AL were similar in magnitude at both ages. There is, however, evidence of situational/method differences that are consistent across age. Genetic influences on actigraph AL in the home were modest (i.e., accounting for less than 30% of variance), whereas shared environmental influences were substantial (i.e., accounting for more than 50% of variance) at both ages. In contrast, actigraph AL in the lab at ages 2 and 3 was highly heritable, accounting for approximately 70% of the total variance, with no significant shared environmental influences. The environments that contributed to variability in actigraph AL in the lab were of the nonshared variety. Both parent and observer ratings of AL showed a similar pattern of high heritability, no shared environmental influences, and significant nonshared influences at each age.

Age-to-age genetic correlations indicating the extent to which genetic effects at age 2 overlap with those at age 3, independent of heritability, show that the genetic factors that

influence actigraph AL in the home at ages 2 and 3 covary perfectly ($r_G=1.0$). No new genetic effects emerged for this measure at age 3. There was also some modest but significant covariance in shared environmental influences across age ($r_C=.28$). Thus, genes and, to a lesser extent, shared environmental influences contribute to continuity in actigraph AL in the home. Change in this measure was due to both shared and nonshared environmental influences. Ninety-two percent of the shared environmental influences ($.51 \div .55 = .92$) and 100% of the nonshared environmental influences on AL at age 3 were independent of prior effects at age 2 (i.e., unique to age 3).

For the remaining measures of AL (actigraph lab, parent and observer ratings), genetic influences contributed to both continuity and change whereas, the environment (nonshared) contributed only to change. The genetic correlations between ages 2 and 3 ranged from .61 for actigraph AL in the lab to .81 for observer-rated AL in the lab, indicating substantial overlap in genetic effects across age. Despite this genetic continuity across age, there was also evidence of significant new genetic variance emerging at age 3. Approximately 63% of the genetic variance of actigraph AL in the lab, 50% of the genetic variance of parent-rated AL, and 34% of the genetic variance of observer-rated AL was unique to age 3.

Discussion

Previous research exploring the factors that influence individual developmental continuity and change in activity level in early childhood have been limited by methodological weaknesses. Additionally, possible situational differences have gone unexplored. This multi-method, multi-situation study addresses these concerns and provides novel evidence of sources of continuity and change in activity level across the transition from infancy to early childhood. The results indicate that the question of genetic and environmental contributions to individual developmental change in activity level is complex and cannot be addressed without considering assessment method or context.

For the actigraph measure of AL in the home, the genetic effects that influenced activity at age 2 were entirely overlapping with those at age 3 indicating that genetic influences contributed to individual continuity not change. Shared environmental influences on AL were substantial at both ages and contributed to *both* continuity and change. Family practices and schedules (e.g., routines, family outings, caretakers' tolerances for activity) are a likely source of common variance between cotwins. To a modest extent, these family-wide influences persist across age, but it is largely genes that explain the phenotypic continuity of AL in the home. Developmental change for this measure of AL was due to environmental factors (both shared and nonshared). The finding of new nonshared environmental effects across age is consistent with prior literature (McGue, Bacon, & Lykken, 1993; Saudino et al., 1996), but novel shared environmental effects at age 3 are particularly interesting as they reflect changes in family-wide experiences perhaps indexing age-related changes in how parents structure their children's environments (e.g., age-appropriate family activities or expectations). These shared environmental influences account for familial similarity in developmental change in AL across early childhood. Nonshared environmental influences explain why the development of AL differs for individuals reared in the same family. That is, the unique child-specific experiences (e.g., accidents, illnesses, and differential treatment) that influence individual differences in AL at age 2 differ from those that influence AL at age 3. Thus, nonshared environmental influences are important to AL in the home in early childhood, but they are age-specific, not enduring.

The actigraph measure of activity in the lab situation displayed a different pattern of results. The phenotypic stability of this activity measure was entirely due to genetic factors, indicating that in this situation, genes, not environments, contribute to continuity across age.

However, developmental change in AL in the lab was due to both nonshared environmental influences and genetic factors. Despite the fact that heritabilities were similar across age, significant new genetic effects on AL in the lab emerged at age 3. Thus, when activity was longitudinally assessed within the context of the lab, genes contributed to individual continuity *and* change. These findings mirror those from MALTS that also found novel genetic variance on activity within the lab at age 3 (Saudino & Cherny, 2001b), however the use of actigraphs in the present study rules-out measure differences as a source of genetic change and allows for more interesting hypotheses regarding the new genetic variance. Because no new genetic variance emerged for actigraph AL in the home, explanations for novel genetic effects at age 3 must focus on the situation. It is possible that the new genetic effects in the lab may reflect children's increased self-control/self-regulation (e.g., Kochanska, Murray, & Harlan, 2000; Kopp, 1982; Rothbart, Ellis, Rueda, & Posner, 2003) or adaptability to novel social situations. This would be consistent with the finding of lower mean levels of activity in the lab at age 3. There is evidence of genetic influences on regulatory temperament dimensions such as inhibitory control/effortful control (e.g., Gagne & Saudino, 2010; Lemery-Chalfant, Doelger, & Goldsmith, 2008), but genetic links between activity level and self-regulation have not yet been explored.

Parent and observer ratings of activity displayed a pattern of results that was similar to the actigraph lab measure (i.e., continuity was due to genetic factors and change was due to both genetic and nonshared environmental influences). The consistency in outcomes between actigraph AL in the lab and observer ratings in the lab is not surprising as previous analyses have found that the two measures are highly correlated and are largely influenced by the same genetic effects (Saudino, 2009). The fact that the results for parent-rated AL mirrored those for the lab-based measures but not the measure of actigraph AL in the home is intriguing. The "home" measure of actigraph AL reflected all activity outside the lab, so broadly speaking, this measure assesses activity in the same context as do the parent ratings. Admittedly, the time frames differ between the two measures (i.e., the TBAQ asks parents about activity behaviors observed within the month, whereas the actigraph data were based on a 2-day data collection period), but the present analyses only included actigraph data for those children whose parents indicated that child behavior during data collection was typical of their usual behavior. Although both measures assess activity in a similar context, prior research has found that they are, for the most part, genetically distinct (Saudino, 2009). Thus, in retrospect, it is not unexpected to find that they display different developmental patterns at the level of etiology. What is interesting, however, is the consistency between both lab measures (i.e., actigraph and observer ratings) and parent ratings in suggesting genetic innovations on activity level at age 3. Perhaps like the lab-based measures of activity, parent ratings of AL on the TBAQ are sensitive to genetically-influenced age-related changes in the child's behavior in novel situations that cannot be fully captured by the 2-day actigraph measures in the home.

The parent rating results in this study differ markedly from those reported in MALTS that found that the phenotypic stability of activity level across early childhood was due to both genetic and nonshared environmental influences (Saudino & Cherny, 2001a). Because in MALTS nonshared environmental contributions to cross-age continuity were found only for parent reports of activity level, and parent ratings on the CCTI measure were found to display contrast effects, it was suggested that the nonshared environmental influences may reflect stability in parent-rating biases that exaggerate differences between twins. The present results support this hypothesis. As expected, parent ratings of activity on the TBAQ did not display twin contrast effects. This likely reflects differences in the types of ratings required on the TBAQ versus the CCTI (i.e., frequency of behaviors versus how accurately an item describes the child's behavior) (Goldsmith & Hewitt, 2003; Hwang & Rothbart, 2003; Saudino, 2003, 2005). Importantly, when contrast effects were not apparent, there was

no evidence of nonshared environmental contributions to the phenotypic stability of parent-rated AL. In MALTS it was suggested that stable contrast biases may explain the higher phenotypic stability of parent ratings as compared to observer ratings. While the MALTS results do suggest that stable rater biases can contribute to the phenotypic stability of activity level, the present results suggest that they cannot fully explain the higher age-to-age correlations for parent ratings as compared to behavioral observations or actigraph measures of AL. That is, in the present study, genetic factors were the *only* source of covariance between parent ratings of activity at ages 2 and 3, yet the stability of parent ratings significantly exceeded that for actigraph AL ($\chi^2=104.84$, 1, $p < .001$). Thus, the higher stability of parent ratings on the TBAQ cannot simply be dismissed as stability in parental rating behaviors.

The finding that genetic influences largely explain the behavioral consistency of activity level in early childhood probably comes as no surprise as it fits with the commonly held static view of genetic effects. Genetic contributions to behavioral consistency are, however, only part of the story. In accordance with a more dynamic view of genetic effects, the present study yields novel evidence indicating that genes also contribute to developmental change in AL across the transition from infancy to early childhood. Genetic sources of change for AL varied depending on method and context of assessment; nonetheless, the overall findings suggest that individual developmental change in AL is in part, a biological process. Although there are genetic innovations across age, the relative importance of genetic influences on individual differences in AL (i.e., heritability) was similar at both ages. This demonstrates that the magnitude of genetic influences on AL does not inform about continuity and change in genetic processes.

For all measures of activity level, environmental factors contributed to individual differences in developmental change. The environmental influences that were most associated with developmental change in AL were of the nonshared variety. The question of which specific nonshared environments influence individual developmental change has generally gone unexplored. Identifying specific nonfamilial environmental factors that influence developmental change remains an important goal for future temperament research.

The limitations of this study should be acknowledged. As is usually the case with developmental research, the sample was predominantly Caucasian, middle-class, and well-educated. Replication with a more diverse sample is needed. However, given that much of what we know about child temperament is based on research with similar sample characteristics the findings have considerable implications. The generalizability of results to a nontwin population is also a question with twin research. It is possible that the substantial shared environmental influences for actigraph AL in the home may reflect twin-specific environments, but this seems unlikely given that a previous study of *nontwin* siblings also showed very high resemblance for a 48-hour measure of actigraph AL (Saudino, Wertz, Gagne, & Chawla, 2004). Finally, the genetic and environmental effects uncovered in twin analyses are anonymous. That is, quantitative genetic analyses indicate the magnitude of genetic influence and the extent of genetic overlap across age, but they do not identify specific genes responsible for individual differences in activity level. Similarly, although these designs can tell us about the impact of shared and nonshared environments on the development of AL, they do not provide information about the specific environments that influence the behaviors under study. Nonetheless, these findings of anonymous effects provide important avenues for future research focusing on specific genetic and environmental effects. For example, for all measures of AL, nonshared environmental influences were an important source of environmental change (the *only* environment source of change for parent-rated AL and both lab-based measures of AL). Researchers interested in understanding how the environment influences developmental change in activity level

would do well to focus on environments that differ *within* families rather than *across* families.

Although not a limitation, there is always the question of links between the temperament dimension of activity level and hyperactivity/ADHD, raising the possibility that the present results could be skewed by the inclusion of children with undiagnosed hyperactivity. Prior analyses with this sample suggest that the genetic influences on activity level do not simply reflect genetic influences on hyperactivity. Multivariate analyses looking at the genetic overlap between actigraph AL and parent reports of hyperactive behavior problems found that the genetic correlation between hyperactivity and actigraph AL was moderate for activity in the home ($r_G=.5$) and modest for activity in the lab ($r_G=.25$) (Iliot, Saudino, Wood, & Asherson, 2010), suggesting that actigraph AL and hyperactivity are to some extent genetically distinct. Moreover, with regard to the present analyses, the findings for actigraph AL in the home and lab were essentially unchanged when the children who were at the top 10% on a measure of hyperactive behavior problems were dropped from the sample. Thus, it does not appear that inclusion of children with emergent hyperactivity has inflated estimates of genetic influence on activity level.

The finding that there are changes in the genetic and environmental factors that operate on individual differences in activity level across age has important implications for developmental researchers interested in AL and other dimensions of temperament. Specifically, findings based on an assessment of temperament at one age may not generalize to another. Similarly, the factors that influence developmental processes in temperament may differ across contexts and/or method of assessment. Clearly, activity level is a complex behavioral dimension and understanding its development will require a multi-method, multi-situational approach. It is likely that this also applies to other temperament dimensions, but this remains an empirical question.

Acknowledgments

The Boston University Twin Project (BUTP) is supported by grant MH062375 to Dr. Saudino from the National Institute of Mental Health. The author gratefully acknowledges the families in the BUTP, the BUTP staff and volunteers. Special thanks to Dr. Jeff Zapfe for his help with the actual data.

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. 2 ed. San Antonio, TX: The Psychological Corporation; 1993.
- Braungart JM, Plomin R, DeFries JC, Fulker D. Genetic influences on tester-rated infant temperament as assessed by Bayley's Infant Behavior Record: Nonadoptive and adoptive siblings and twins. *Developmental Psychology*. 1992; 28:40–47.
- Brody G, Stoneman Z, Burke M. Child temperaments, maternal differential behavior, and sibling relationships. *Developmental Psychology*. 1987; 23:354–362.
- 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]
- Cyphers LH, Phillips K, Fulker DW, Mrazek DA. Twin temperament during the transition from infancy to early childhood. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1990; 29:392–397. [PubMed: 2347836]

- Deater-Deckard K, Petrill SA, Thompson LA. Anger/frustration, task persistence, and conduct problems in childhood: a behavioral genetic analysis. *Journal of Child Psychology and Psychiatry*. 2007; 48(1):80–87. [PubMed: 17244273]
- Deater-Deckard K, Petrill SA, Thompson LA, DeThorne LS. A cross-sectional behavioral genetic analysis of task persistence in the transition to middle childhood. *Developmental Science*. 2005; 8(3):F21–F26. [PubMed: 15819750]
- Deater-Deckard K, Petrill SA, Thompson LA, DeThorne LS. A longitudinal behavioral genetic analysis of task persistence. *Developmental Science*. 2006; 9(5):498–504. [PubMed: 16911452]
- Eaton WO, Enns LR. Sex differences in human motor activity level. *Psychological Bulletin*. 1986; 100:19–28. [PubMed: 3526375]
- Eaton WO, McKeen N, Campbell D. The waxing and waning of movement: Implications for psychological development. *Developmental Review*. 2001; 21:205–223.
- Eaton WO, McKeen N, Lam C. Instrumented motor activity measurement of the young infant in the home: validity and reliability. *Infant Behavior & Development*. 1988; 11:375–378.
- Eaton, WO.; McKeen, N.; Saudino, KJ. Measuring human individual differences in general motor activity with actometers. In: Ossenkopp, K-P.; Kavaliers, M.; Sanberg, PR., editors. *Techniques in the behavioral and neural sciences: motor activity and movement disorders*. Amsterdam: Elsevier; 1996. p. 79-92.
- 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]
- Emde, RN.; Hewitt, JK., editors. *The Transition from Infancy to Early Childhood: Genetic and Environmental Influences in the MacArthur Longitudinal Twin Study*. New York: Oxford University Press; 2001.
- Fagan J. The interaction between child sex and temperament in predicting behavior problems of preschool-age children in day care. *Early Child Development and 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.
- Finn KJ, Specker B. Comparison of Acitwatch activity monitor and Children's Activity Rating Scale in children. *Medicine & Science in Sports & Exercise*. 2000; 32:1794–1797. [PubMed: 11039655]
- Freedson, PS.; Sirard, J.; Debold, N.; Pate, R.; Dowda, M.; Sallis, J. Validity of two physical activity monitors in children and adolescents; Paper presented at the XIXth Symposium of European Group of Pediatric Work Physiology; 1997.
- Gagne JR, Saudino KJ. Wait for it! A twin study of inhibitory control in early childhood. *Behavior Genetics*. 2010; 40:327–337. [PubMed: 19936910]
- Gjone H, Stevenson J. A longitudinal twin study of temperament and behaviour problems: common genetic or environmental influences? *Journal of the American Academy of Child & Adolescent Psychiatry*. 1997; 36:1448–1456. [PubMed: 9334559]
- Goldsmith HH. Studying temperament via construction of the Toddler Behavior Assessment Questionnaire. *Child Development*. 1996; 67:218–235. [PubMed: 8605830]
- Goldsmith HH, Buss A, Plomin R, Rothbart MK, Thomas A, Chess S, et al. Roundtable: What is temperament? Four approaches. *Child Development*. 1987; 58:505–529. [PubMed: 3829791]
- Goldsmith HH, Buss KA, Lemery KS. Toddler and childhood temperament: Expanded content, stronger genetic evidence, new evidence for the importance of environment. *Developmental Psychology*. 1997; 33:891–905. [PubMed: 9383612]
- Goldsmith HH, Hewitt EC. Validity of parental report of temperament: Distinctions and needed research. *Infant Behavior and Development*. 2003; 26:108–111.
- 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.
- Guralnick MJ, Groom JM. The correspondence between temperament and peer interactions for normally developing and mildly delayed preschool children. *Child: Care, Health and Development*. 1990; 16:165–175.

- Hall S, Halperin JM, Schwartz S, Newcorn J. Behavioral and executive functions in children with attention-deficit hyperactivity disorder and reading disability. *Journal of Attention Disorders*. 1997; 1:235–247.
- Heil DP. Predicting Activity Energy Expenditure Using the Actical Activity Monitor. *Research Quarterly for Exercise and Sport*. 2006; 77:64–80.
- Hollingshead, A. Four-factor Index of Social Status. Yale University; 1975.
- Hwang J, Rothbart MK. Behavior genetics studies of infant temperament: Findings vary across parent-report instruments. *Infant Behavior and Development*. 2003; 26:112–114.
- Ilott N, Saudino KJ, Wood A, Asherson P. A genetic study of ADHD and activity level in infancy. *Genes Brain and Behavior*. 2010; 9(3):296–304.
- Jewsuwan R, Luster T, Kostelnik M. The relation between parents' perceptions of temperament and children's adjustment to preschool. *Early Childhood Research Quarterly*. 1993; 8:33–51.
- Katainen S, Räeikköenen K, Keltikangas-Järvinen L. Childhood temperament and mother's child-rearing attitudes: Stability and interaction in a three-year followup study. *European Journal of Personality*. 1997; 11:249–265.
- Katainen S, Räeikköenen K, Keltikangas-Järvinen L. Development of temperament: Childhood temperament and the mother's childrearing attitudes as predictors of adolescent temperament in a 9-yr follow-up study. *Journal of Research on Adolescence*. 1998; 8:485–509.
- Klein HA, Ballantine JH. The relationship of temperament to adjustment in British infant schools. *Journal of Social Psychology*. 1988; 128:585–595. [PubMed: 3210671]
- Kochanska G, Murray KT, Harlan ET. Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology*. 2000; 36(2): 220–232. [PubMed: 10749079]
- Konrad K, Gunther T, Heinzel-Gutenbrunner M, Herpertz-Dahlmann B. Clinical evaluation of subjective and objective changes in motor activity and attention in children with attention-deficit/hyper-activity disorder in a double-blind methylphenidate trial. *Journal of Child and Adolescent Psychopharmacology*. 2005; 15:180. [PubMed: 15910203]
- Kopp CB. Antecedents of self-regulation: A developmental perspective. *Developmental Psychology*. 1982; 18(2):199–214.
- Laible D, Panfile T, Makariev D. The Quality and Frequency of Mother-Toddler Conflict: Links With Attachment and Temperament. *Child Development*. 2008; 79(2):426–443. [PubMed: 18366432]
- Lemery-Chalfant K, Doelger L, Goldsmith HH. Genetic relations between effortful and attentional control and symptoms of psychopathology in middle childhood. *Infant and Child Development*. 2008; 17(4):365–385.
- Matheny AP. Bayley's infant behavior record: Behavioral components and twin analyses. *Child Development*. 1980; 51:1157–1167. [PubMed: 7193557]
- Matheny AP. Children's behavioral inhibition over age and across situations: genetic similarity for a trait during change. *Journal of Personality*. 1989; 57:215–235. [PubMed: 2769555]
- Matier-Sharma K, Perachio N, Newcorn JH, Sharma V, Halperin JM. Differential diagnosis of ADHD: are objective measures of attention, impulsivity, and activity level helpful? *Child Neuropsychology*. 1995; 1:118–127.
- McGrath AM, Handwerk ML, Armstrong KJ, Lucas CP, Friman PC. The validity of the ADHD section of the Diagnostic Interview Schedule for Children. *Behavior Modification*. 2004; 28:349. [PubMed: 15104867]
- McGue M, Bacon S, Lykken DT. Personality stability and change in early adulthood: a behavioural genetic analysis. *Developmental Psychology*. 1993; 29:96–109.
- McGue M, Bouchard TJ. Adjustment of twin data for the effects of age and sex. *Behavior Genetics*. 1984; 14:325–343. [PubMed: 6542356]
- Neale, MC.; Boker, SM.; Xie, G.; Maes, HH. *Mx: Statistical modeling*. 7th Edition. Richmond, VA: Virginia Commonwealth University; 2006.
- Neale, MC.; Cardon, LR. *Methodology for genetic studies of twins and families*. Dordrecht, the Netherlands: Kluwer Academic Publishers; 1992.

- O'Donnell JP, Van Tuinan M. Behavior problems of preschool children: dimensions and congenital correlates. *Journal of Abnormal Child Psychology*. 1979; 7:61–75. [PubMed: 438432]
- Olson SL, Bates JE, Sandy JM, Lanthier R. Early developmental precursors of externalizing behavior in middle childhood and adolescence. *Journal of Abnormal Child Psychology*. 2000; 28:119–133. [PubMed: 10834765]
- 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]
- Plomin, R. Developmental behavioral genetics: Stability and instability. In: Bornstein, MH.; Krasnegor, NA., editors. *Stability and continuity in mental development*. Hillsdale, NJ: Erlbaum; 1989.
- Plomin R, Campos JJ, Corley R, Emde RN, Fulker DW, Kagan J, et al. Individual differences during the second year of life: The MacArthur Longitudinal Twin Study. Individual differences in infancy: Reliability, stability, and predictability. 1990:431–455.
- Plomin, R.; DeFries, JC. *Origins of individual Differences in linfancy: The Colorado Adoption Project*. Toronto: Academic Press; 1985.
- Plomin R, Emde RN, Braungart JM, Campos J, Corley R, Fulker DW. Genetic change and continuity from 14 to 20 months: The MacArthur Longitudinal Twin Study. *Child Development*. 1993; 64:1354–1376. [PubMed: 8222877]
- Plomin R, Nesselroade JR. Behavioural genetics and personality changes. *Journal of Personality*. 1990; 58:191–220. [PubMed: 2198339]
- Posthuma D, Beem AL, Eco JC, van Baal GCM, von Hjelmberg JB, Iachine I, et al. Theory and Practice in Quantitative Genetics. *Twin Research*. 2003; 6(5):361–376. [PubMed: 14624720]
- 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]
- Roberts BW, DelVecchio WF. The rank-order consistency of personality traits from childhood to old age: A quantiative review of longitudinal studies. *Psychological Bulletin*. 2000; 126(1):3–25. [PubMed: 10668348]
- Rothbart MK. Longitudinal observation of infant temperament. *Developmental Psychology*. 1986; 22:356–365.
- Rothbart MK, Ellis LK, Rueda MR, Posner MI. Developing mechanisms of temperamental effortful control. *Journal of Personality*. 2003; 71(6):1113–1143. [PubMed: 14633060]
- Saudino KJ. Parent ratings of infant temperament lessons from twin studies. *Infant Behavior & Development*. 2003; 26:100–107.
- Saudino KJ. Behavioral genetics and child temperament. *Journal of Developmental and Behavioral Pediatrics*. 2005; 26(3):1.
- Saudino KJ. Do Different Measures Tap the Same Genetic Influences? A Multi-method Study of Activity Level in Young Twins. *Developmental Science*. 2009; 12:626–633. [PubMed: 19635088]
- Saudino, KJ.; Cherny, SS. Parent ratings of temperament in twins. In: Emde, RN.; Hewitt, JK., editors. *The transition from infancy to early childhood: Genetic and environmental influences in the MacArthur Longitudinal Twin Study*. NY: Oxford University Press; 2001a. p. 73-88.
- Saudino, KJ.; Cherny, SS. Sources of continuity and change in observed temperament. In: Emde, RN.; Hewitt, JK., editors. *The transition from infancy to early childhood: Genetic and environmental influences in the Mac Arthur Longitudinal Twin Study*. NY: Oxford University Press; 2001b. p. 89-110.
- Saudino KJ, Cherny SS, Plomin R. Parent ratings of temperament in twins: Explaining the "too low" DZ correlations. *Twin Research*. 2000; 3:224–233. [PubMed: 11463143]
- Saudino KJ, Eaton WO. Infant temperament and genetics: An objective twin study of motor activity level. *Child Development*. 1991; 62:1167–1174. [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; 13:81–95.
- Saudino KJ, McGuire S, Reiss D, Hetherington EM, Plomin R. Parent ratings of EAS temperaments in twins, full siblings, half siblings, and step siblings. *Journal of Personality and Social Psychology*. 1995; 68:723–733. [PubMed: 7738774]

- 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, 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*. 2008; 79:930–943. [PubMed: 18717899]
- Schaughency EC, Fagot BI. The prediction of adjustment at age 7 from activity level at age 5. *Journal of Abnormal Child Psychology*. 1993; 21(1):29–50. [PubMed: 8463503]
- Singer JD. Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*. 1998; 24(4):322–355.
- Stevenson J, Fielding J. Ratings of temperament in families of young twins. *British Journal of Developmental Psychology*. 1985; 3:143–152.
- Teglasi H, MacMahon BH. Temperament and common problem behaviors of children. *Journal of Applied Developmental Psychology*. 1990; 11:331–349.
- Wilson, RS.; Matheny, AP. Behavior genetics research in infant temperament: The Louisville Twin Study. In: Plomin, R.; Dunn, J., editors. *The study of temperament: Changes, continuities and challenges*. Hillsdale, NJ: Erlbaum; 1986. p. 81-97.
- Wohlwill, JF. *The study of behavioral development*. Oxford, UK: Academic Press; 1973.
- 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]
- Zajdeman HS, Minnes PM. Predictors of children's adjustment to day-care. *Early Child Development and Care*. 1991; 74:11–28.

Figure 1.

Longitudinal Model. AL=Activity level. A_2 , C_2 , and E_2 = genetic, shared environmental and nonshared environmental effects that are common to ages 2 and 3; A_3 , C_3 , and E_3 = genetic and environmental effects that are unique to age 3. a_{21} , a_{22} , a_{31} = genetic path co-efficients. s_2 and s_3 = sibling interaction parameters at ages 2 and 3, respectively (included only in parent rating model). MZ=monozygotic twins, DZ=dizygotic twins.

Table 1

Means, Standard Deviations and Tests of Age and Sex Effects for AL Measures

AL Measure	Age 2		Age 3		Age	Sex	Age×Sex
	Males M (SD)	Females M (SD)	Males M (SD)	Females M (SD)	F(df)	F(df)	F(df)
Actigraph Home ^a	463.69 (106.92)	438.68 (95.18)	512.30 (108.55)	494.10 (105.33)	33.50** (1, 314)	4.36* (1, 328)	0.54 (1, 319)
Actigraph Lab ^a	739.42 (276.67)	602.64 (225.02)	626.58 (230.09)	587.67 (196.69)	45.23** (1, 314)	29.35** (1, 320)	17.47** (1, 313)
Parent Rating	43.40 (9.27)	39.80 (7.56)	41.58 (9.28)	37.39 (8.24)	13.10** (1, 299)	16.76** (1, 317)	0.50 (1, 299)
Observer Rating Lab	0.20 (0.81)	-0.22 (0.74)	0.11 (0.86)	-0.13 (0.76)	3.07† (1, 306)	30.74** (1, 318)	5.08* (1, 305)

Note. AL=Activity level.

^a Average Actical activity count per minute.

† $p < .10$.

* $p < .05$.

** $p < .01$.

Phenotypic Stability of Activity Level and Intercorrelations Between Activity Measures at Ages 2 and 3

Table 2

	Age-to-Age <i>r</i>	Actigraph Home	Actigraph Lab	Parent Rating	Observer Rating Lab
Actigraph Home	.43**	1	.40**	.25**	.18**
Actigraph Lab	.46**	.34**	1	.16*	.68**
Parent Rating	.62**	.19**	.18**	1	.19**
Observer Rating Lab	.45**	.17**	.69**	.15**	1

Note. Correlations above the diagonal are between measures at age 2, below the diagonal are between measures at age 3.

*
p < .05.
**
p < .01.

Table 3

Twin Intraclass Correlations, Cross-Age Correlations

AL Measure	Intraclass Correlations		Cross-Age Correlations	
	MZ (<i>n</i> pairs)	DZ (<i>n</i> pairs)	MZ (<i>n</i> pairs)	DZ (<i>n</i> pairs)
Actigraph Home				
Age 2	.87** (143)	.70** (166)		
Age 3	.82** (140)	.67** (161)	.50** (139)	.23** (161)
Actigraph Lab				
Age 2	.74** (143)	.36** (167)		
Age 3	.72** (140)	.41** (161)	.49** (139)	.25** (161)
Parent Rating				
Age 2	.82** (142)	.46** (163)		
Age 3	.82** (134)	.53** (157)	.58** (132)	.37** (156)
Observer Rating Lab				
Age 2	.54** (144)	.38** (167)		
Age 3	.59** (140)	.38** (163)	.48** (139)	.28** (162)

Note. AL=Activity level, MZ=monozygotic twins, DZ=dizygotic twins.

*
 $p < .05$.

**
 $p < .01$.

Table 4

Fit Statistics for Longitudinal Cholesky Model

Variable and Model	Overall Fit of Model ^a					Relative Fit of Model ^b		
	-2LL	df	χ^2	df	p	AIC	$\Delta\chi^2$	df
Actigraph Home								
Saturated	3971.25	1202						
ACE	3983.92	1213	12.67	11	.32	-9.33		
CE	4021.23	1216	49.98	14	.00	12.98	37.31	3
AE	4051.56	1216	80.31	14	.00	52.31	67.64	3
E	4537.01	1219	565.76	17	.00	531.76	553.08	6
Actigraph Lab								
Saturated	3950.19	1205						
ACE	3968.95	1216	18.76	11	.07	-3.24		
CE	4005.86	1219	55.67	14	.00	27.67	36.91	3
AE	3972.14	1219	21.95	14	.08	-6.05	3.19	3
E	4212.72	1222	262.53	17	.00	228.53	243.77	6
Parent Rating								
Saturated	2441.42	1180						
ACEs	2454.32	1189	12.90	9	.17	-5.10		
ACE	2455.93	1191	14.51	11	.21	-7.49	1.61 ^c	2
CE	2541.58	1194	100.16	14	.00	72.16	85.65	3
AE	2458.47	1194	17.05	14	.25	-10.95	2.54	3
E	2802.19	1197	360.77	17	.00	326.77	346.26	6
Observer Rating Lab								
Saturated	2651.63	1208						
ACE	2658.48	1219	6.86	11	.81	-15.14		
CE	2671.32	1222	19.70	14	.14	-8.30	12.84	3
AE	2663.34	1222	11.72	14	.63	-16.29	4.86	3
E	2813.00	1225	162.37	17	.00	128.37	155.51	6

Note. -2LL=Likelihood Statistic, χ^2 =Chi-square fit statistic, df=Degrees of freedom, AIC=Akaike's Information Criterion.

^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.

^c Relative fit of the model determined by the $\Delta\chi^2$ between the ACEs (sibling interaction model) and the full ACE model. Best fitting model indicated in bold.

Table 5

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

AL Measure	Variance Components									Covariances			
	Age 2			Age 3			Change						
	a^2	c^2	e^2	a^2	Continuity					e^2	a^2	c^2	e^2
					c^2	e^2							
Actigraph Home	.29 (.16-.44)	.58 (.43-.69)	.14 (.11-.18)	.25 (.10-.42)	.04 (.01-.13)	.00 (.00-.01)	.00 (-.18-.18)	.51 (.36-.60)	.20 (.15-.25)	1.0 (.65-1.0)	.28 (.04-.47)	-.03 (-.17-.25)	
Actigraph Lab	.72 (.65-.78)	--	.28 (.22-.35)	.25 (.17-.35)	--	.00 (.00-.01)	.42 (.32-.51)	--	.32 (.26-.40)	.61 (.50-.77)	--	.04 (-.11-.18)	
Parent Rating	.84 (.79-.87)	--	.16 (.13-.21)	.43 (.34-.51)	--	.00 (.00-.01)	.41 (.33-.49)	--	.16 (.13-.21)	.71 (.64-.78)	--	.09 (-.21-.24)	
Observer Rating	.55 (.45-.64)	--	.45 (.36-.55)	.39 (.27-.52)	--	.00 (.00-.02)	.20 (.05-.33)	--	.42 (.33-.51)	.81 (.68-.96)	--	-.08 (-.22-.06)	

Note. AL=Activity level, 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. Dashes indicate that the parameter was not included in the best-fitting model.