HHS Public Access

Author manuscript

Dev Psychopathol. Author manuscript; available in PMC 2022 August 01.

Published in final edited form as:

Dev Psychopathol. 2022 February ; 34(1): 321–333. doi:10.1017/S0954579420001017.

How nonshared environmental factors come to correlate with heredity

Christopher R. Beam[1], **Patrizia Pezzoli**[2], **Jane Mendle**[3], **S. Alexandra Burt**[4], **Michael C. Neale**[5], **Steven M. Boker**[6], **Pamela K. Keel**[7], **Kelly L. Klump**[4]

[1]Department of Psychology, University of Southern California, Los Angeles, CA, USA

[2]Department of Psychology, Åbo Akademi University, Abo, Finland

[3]Department of Human Development, Cornell University, Ithaca, NY, USA

[4]Department of Psychology, Michigan State University, East Lansing, MI, USA

[5]Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA

[6]Department of Psychology, University of Virginia, Charlottesville, VA, USA

[7]Department of Psychology, Florida State University, Tallahassee, FL, USA

Abstract

Conventional longitudinal behavioral genetic models estimate the relative contribution of genetic and environmental factors to stability and change of traits and behaviors. Longitudinal models rarely explain the processes that generate observed differences between genetically and socially related individuals. We propose that exchanges between individuals and their environments (i.e., phenotype–environment effects) can explain the emergence of observed differences over time. Phenotype–environment models, however, would require violation of the independence assumption of standard behavioral genetic models; that is, uncorrelated genetic and environmental factors. We review how specification of phenotype–environment effects contributes to understanding observed changes in genetic variability over time and longitudinal correlations among nonshared environmental factors. We then provide an example using 30 days of positive and negative affect scores from an all-female sample of twins. Results demonstrate that the phenotype–environment effects explain how heritability estimates fluctuate as well as how nonshared environmental factors persist over time. We discuss possible mechanisms underlying change in gene–environment correlation over time, the advantages and challenges of including gene–environment correlation in longitudinal twin models, and recommendations for future research.

Author for Correspondence: Christopher R. Beam, Assistant Professor of Psychology and Gerontology, Department of Psychology, University of Southern California, 3620 McClintock Ave, Seeley G. Mudd, Room 501, Los Angeles, CA 90089; beamc@usc.edu.

Supplementary Material. The supplementary material for this article can be found at https://doi.org/10.1017/S0954579420001017.

Conflicts of Interest. None

Keywords

affect; developmental behavioral genetics; gene–environment interplay; longitudinal modeling; mood

In psychopathology research, developmental behavioral genetic studies quantify relative contributions of heritable and environmental factors to phenotypic variation but tend to remain agnostic to underlying causal processes. The great majority of longitudinal twin studies show that heritable factors account for stability of phenotypes, such as traits, behaviors, and emotions, whereas environmental factors account for change (Bartels et al., 2004; Bartels, Rietveld, Van Baal, & Boomsma, 2002; Briley et al., 2019; Bronfenbrenner & Ceci, 1994; Eaves, Long & Health, 1986). These studies are largely descriptive, and do not address causal processes that explain accrual of differences in phenotypic outcomes over time (for exceptions, see Dolan, De Kort, Van Beijsterveldt, Bartels, & Boomsma, 2014; Neale & McArdle, 2000; van den Berg, Beem, & Boomsma, 2006). One limitation of the often-used additive genetic models is that they rely on the assumption that genetic and environmental factors are uncorrelated (Polderman et al., 2015). As, in reality, genes and environments do correlate and interact with one another, developmental behavioral genetic models should incorporate this interdependence. While gene–environment interplay is accounted for in models that assume independence of unmeasured genetic and environmental components (e.g., Johnson, 2007; Neale & Cardon, 1992; Scarr & McCartney, 1983), the covariance between latent genetic and environmental factors is seldom incorporated into developmental behavioral genetic models. Doing so would improve our ability to investigate the processes that contribute to variability in human complex traits.

The present study is not the first call to address this limitation. In 1958, Anastasi commented, “as we proceed along the continuum of indirectness, the range of variation of possible outcomes of hereditary factors expands rapidly. At each step in the causal chain, there is fresh opportunity for interaction with other hereditary factors as well as with environmental factors” (Anastasi, 1958, p. 199). The conventional assumption that genetic and environmental factors are uncorrelated breaks an important link in the causal chain of Anastasi’s continuum of indirectness: how people interface with their environments. Individuals receive, select, and evoke responses from their environments in systematic ways. Known as gene–environment correlation (*r*GE), nonrandom exposure to environments based on genetically influenced characteristics is part of what it means to be a person embedded within continuously changing social environments (Eaves, Krystyna, Martin, & Jinks, 1977; Plomin, DeFries, & Loehlin, 1977; Scarr & McCartney, 1983). This phenomenon arises from various processes (Plomin et al., 1977). Passive *r*GE occurs when both genes and environments are provided to individuals (e.g., parents pass along genes for affective dysregulation and provide chaotic home environments). Active (or selective) *r*GE occurs when individuals nonrandomly select environments syntonic with their genetically influenced characteristics (e.g., persons with high risk of affective dysregulation select affectively dysregulated peers). Evocative *r*GE occurs when environments nonrandomly reinforce genetically influenced characteristics (e.g., persons with high risk of affective

dysregulation elicit greater interpersonal conflict from their environments). Nonetheless, *r*GE is conventionally unaccounted for in developmental and, more generally, longitudinal twin studies. We note that longitudinal co-twin control studies, however, do adjust for *r*GE while examining how nonshared environmental factors contribute to development (McGue, Osler, & Christensen, 2010; Røysamb & Tambs, 2016). The current aim is to modify longitudinal twin models to incorporate *r*GE processes that represent real-time processes (Briley et al., 2019). By doing so, we improve our understanding of the role of *r*GE in accounting for various phenotypic processes over time.

Longitudinal twin models miss an important feature of the causal chain between genotype and environments: namely, persons. As argued by Scarr and McCartney, "some genotypes are more likely to receive and select certain environments than others" (1983, p. 426). Yet *genotypes* do not receive and select environments; people do (Turkheimer & Waldron, 2000). Numerous intermediate and bidirectional processes connect genetic and environmental predispositions with complex traits (Cole, 2009; Gottlieb, 2003). Intermediary pathways that connect genetic factors to environments include, among others, cognition, personality, and affect (Cole, 2009; Gottlieb, 2003). In particular, genetic influences underlying affective dysregulation are mediated by processes "external" to individuals, such as changes in social environments, as well as by processes "internal" to individuals, such as hormones (e.g., cortisol) and other molecular processes (e.g., methylation activity affecting RNA transcription). Consequently, individuals are a critical node in the processes underlying *r*GE.

Reciprocal Effects Models

Among the processes that best exemplify Phenotype-Environment (P→E) associations is within-family diversification, or "sibling drift". Several efforts have been made to study this phenomenon by incorporating *r*GE in longitudinal twin models (e.g., Beam & Turkheimer, 2013; de Kort, Dolan, & Boomsma, 2012; Dolan et al., 2014; Neale & Cardon, 1992). In particular, reciprocal effects models (REMs) have guided our thinking about change processes (Bell, 1968; Scarr, 1992; Winship & Korenman, 1999; Beam & Turkheimer, 2017). REMs summarize how small phenotypic advantages (e.g., calm demeanor) translate into superior abilities (e.g., affective stability). Individuals are initially provided with supportive environments (e.g., quiet vs. chaotic household), which in turn predispose them to seek out novel correlated environments (e.g., organized vs. disorganized peers). Further reinforcement of phenotypic advantages in these novel environments may then lead to the provision of additional support for affective stability (e.g., adults are more likely to mentor and model coping strategies). Over time, phenotypes become increasingly correlated with environments via iterative person–environment matching. To the extent that phenotypes are genetically influenced, genetic influences will themselves become increasingly correlated with environments with time.

In previous work, Dickens and Flynn's (2001) version of REMs has been invoked to explain how fleeting exogenous environmental experiences could explain time-limited rank order differences in cognitive ability between genetically related individuals (Dickens, Turkheimer, & Beam, 2011). REMs subsequently were adapted to explain how P→E

matching processes might cause genetically related individuals to drift apart over time (Beam & Turkheimer, 2013; Beam et al., 2015, 2016). Importantly, we showed that the degree of genetic relatedness between individuals influences the degree of similarity and difference between their environments, which then underlies how different individuals eventually become. In other words, the smaller the genetic relatedness between individuals, the greater they will drift apart over time.

Figure 1 shows a multilevel REM, referred to here as the P→E model, in which total genetic effects are divided into genetic and environmental effects shared by twins in the same family (A^b and E^b, respectively) and genetic and environmental effects unshared by twins in the same family (A^w and E^w, respectively). At the within-level only, there is an autoregressive parameter (b_{PE}) projecting from individuals' phenotypic scores at time t (P_t) to individuals' unshared (nonshared) environments at time t+1 (E^w_{t+1}). Although similar to a modified genetic simplex model, where genetic and environmental factors unidirectionally influence phenotypes across repeated measurements (Boomsma & Molenaar, 1987; Eaves et al., 1986), this autoregressive parameter represents the influence individuals have on their future environments. The autoregressive parameter generates within-family *r*GE, that is, person–specific correlations between genotype and environments. Adding this parameter produces different model expectations for monozygotic (MZ) twins and dizygotic (DZ) twins. Moreover, it more accurately models how nongenetically identical siblings differentiate over time compared to MZ twins. According to path tracing rules (Boker & McArdle, 2014), the autoregressive parameter necessarily induces accumulation of *r*GE across time as well as correlations among nonshared environmental components (de Kort et al., 2012; Eaves et al., 1977).

Traditional longitudinal twin models adhere to the assumption of independence between genetic and environmental factors. As a result, they generate upwardly biased genetic variance estimates (Beam & Turkheimer, 2013). In other words, such models might erroneously conclude that heritability increases when it does not (Purcell, 2002). At the same time, accumulation of unmodeled *r*GE in genetic variance components might lead to downwardly biased longitudinal correlations among nonshared environmental components. Traditional longitudinal twin studies tend to report uncorrelated structures of nonshared environmental factors over time, that is, age-specific influences on phenotypic outcomes (for examples, see Bartels et al., 2004 [developmental psychopathology]; Klump, Burt, Mcgue, & Iacono, 2007 [eating disorders]; Nivard et al., 2015 [depressive symptomatology]; Petrill et al., 2004 [cognitive ability]). This is justified with the notion that, within families, there is no rank-order stability between twins over time, so that twins are re-sorted randomly as time passes. However, since individuals are nonrandomly exposed to environments due to genetic and environmental reasons (Plomin et al., 1977; Scarr & McCartney, 1983), this notion appears unrealistic. It is unlikely, for instance, that the more temperamental of two siblings would suddenly become the steadier sibling, without ongoing selection and reinforcement of environments that stabilize affect. As life plods along, environmental influences stabilize, suggesting that twins' environments actually become canalized (Dickens et al., 2011). Given that genes and environments do correlate, P→E models might provide a more accurate representation of longitudinal processes. Accordingly, the P→E model depicted in Figure

1 accommodates violations of the independence assumption, thereby permitting *r*GE to accumulate over time.

As an example, affect is genetically and environmentally influenced (Montag et al., 2016), and this has implications for how individuals engage with their social environments. Two siblings – one more affectively labile than the other – are likely to select and evoke different environmental milieus. Compared to the less affectively labile sibling, the more labile sibling may be drawn to individuals who are more emotionally expressive and engage in more argumentative and baiting behaviors. As a result, this more affectively labile sibling may experience and contribute to a social environment characterized by more conflict, leading to still greater labile affect over time. However, while we have already demonstrated that small within-family phenotypic differences put genetically related siblings onto different cognitive (Beam et al., 2015) and personality (Beam & Sharp, 2020) trajectories, no previous study has applied P→E models to affect.

One consequence of model estimated *r*GE is that the interpretation of nonshared environmental components changes (Beam, Turkheimer, Dickens, & Davis, 2015; de Kort et al., 2012). In P→E models, each nonshared environmental component consists of two components: a part that correlates with genotype (i.e., it differs due to genotypic differences) and a part that is uncorrelated with genotype (i.e., occasion-specific residuals). In the former, genetic factors indirectly influence nonshared environmental factors, by definition, and are therefore confounded with genotype. Comparisons of the longitudinal structure of nonshared environmental components between models that accommodate *r*GE and models that assume independence between genetic and environmental factors provide a straightforward empirical approach to determining whether accommodating *r*GE alters the meaning of the nonshared environment. In one previous report, we have shown that differences in nonshared environmental structures between models that accommodate *r*GE and models that do not are greater in DZ groups than MZ groups, suggesting that accommodation of *r*GE helps to explain why environmental factors might be more correlated over time (Beam et al., 2015).

Quality of Environments and Measurement Density

Two conditions affect estimation of within-family *r*GE but have not yet been considered: the assumption that individuals have a wide range of environments available to them and the intervals of time between measurements (e.g., days vs. weeks vs. years). First, as posited by Fuller and Thompson (1960), enriched environments allow individuals to utilize environments freely, which support genetic expression of traits and allow individuals to adapt differently from one another (Bronfenbrenner & Ceci, 1994; Fischbein, 1978). In contrast, restricted environments minimize behavioral options, leading individuals to behave more similarly, irrespective of genotypic differences. Therefore, siblings who are less genetically related (e.g., fraternal twins, full siblings, half-siblings, and so forth), and who have the resources to select from a range of environments, are expected to become less similar to each other over time. Instead, genetically identical individuals are more likely to select similar environments and, in turn, to become more similar over time. Statistically, this would imply that heritability increases over time, owing to the fact that within-pair MZ correlations are predicted to remain stable over time while non-MZ correlations are expected

to decline. One implication of these expectations is that within-family *r*GE ought to increase over time, as genetically dissimilar sibling pairs develop. As within-family *r*GE can be tested statistically, these competing hypotheses can be evaluated directly.

Second, intervals of time between measurements can affect estimates of within-family *r*GE. As intervals of time between measurements increase, estimated autocorrelation between measurement occasions will decrease, as long as the process is stationary and noncyclic. Because longitudinal twin studies are laborious and expensive, dense measurement schedules of phenotypes are rare, and the temporal dynamics of *r*GE remain poorly understood. Shorter intervals between measurements, such as daily measurement schedules, could lead to greater estimates of the magnitude of the within-family *r*GE compared to longer intervals between measurements, such as annual measurement schedules. Developmentally, however, shorter divisions of time might not allow enough time for persons to receive, select, and evoke environments critical enough to make a difference for certain phenotypes, especially in view of evidence that effects of twins' unique measured environments on behavior are small (Turkheimer & Waldron, 2000). Thus, to best understand how environments guide change in genetically influenced characteristics, intervals of measurement should be suitably matched to expected change in phenotypes of interest.

The Current Study

In the present study, we explored the temporal dynamics of *r*GE using daily measurements of affect in a young all-female sample of twin pairs aged 16–25. Specifically, we investigated how nonrandom matching between these participants and their environments accounted for within-family differences in affect over 30 days. While some traits change slowly over a long period of time (e.g., personality, see Roberts & Mroczek, 2008), affect has been shown to fluctuate daily (Röcke, Li, & Smith, 2009). As a result, genetic and environmental influences may, too, fluctuate. Therefore, the 30-day study window is a circumscribed period of time in which we can explore whether genetic and environmental influences underlying affect fluctuate. Using a P→E model, we tested two hypotheses: (a) heritability tends to be over-estimated for daily affect when *r*GE is unmodeled; *and* (b) within-family *r*GE will generate greater correlations among nonshared environmental factors over time. With respect to the latter hypothesis, we further expected environmental factors more proximally situated in time to be more correlated than environmental factors less proximally situated in time. Both hypotheses would suggest that the extent to which individuals engage with their environments on a daily basis, based on genotypically influenced characteristics, affects their nonrandom selection of environments, causing them to gravitate toward more like environments from day-to-day.

Of note, as the sample consists only of young women, changes across the menstrual cycle in the levels of two ovarian steroid hormones in the female body, estrogen and progesterone, may also account for differences in daily affect. Whereas estrogen is the predominant hormone produced in the first, proliferative half of the cycle (follicular phase), progesterone predominates in the secretory half of the cycle following ovulation (luteal phase), albeit estrogen also remains somewhat elevated (Poromaa & Gingnell, 2014). This

cyclical fluctuation in neuroendocrine levels has consequences including variation in affect (Farage, Neill, & MacLean, 2009; Romans, Clarkson, Einstein, Petrovic, & Stewart, 2012). The luteal phase, for example, may engender a high-risk phase for increased negative and depressed affect as a result of higher progesterone levels (Bäckström et al., 2011; Halbreich et al., 2012). Although changes in affect across the menstrual cycle are not the primary focus of the current study, data collected over the menstrual cycle raise two pertinent issues. First, the cyclic nature of estrogen and progesterone implies that the correlation between individuals and environments is non-stationary and may vary in an oscillatory manner. Second, estrogen and progesterone direct genomic actions (e.g., gene transcription) and have been found to moderate genetic influences underlying behavioral phenotypes, such as binge eating, independent of genetic influences underlying ovarian hormones (Klump et al., 2015). Estrogen and progesterone, thus, may interact with genotype causing heritability to fluctuate, possibly in addition to effects of *r*GE via hypothesized person–environment matching. Therefore, to the extent that ovarian hormones act upon genetic components that differ within twins, they constitute an additional "environmental" source of variance in the P→E system (Purcell, 2002), as our model cannot distinguish between endogenous and exogenous environmental factors that vary within twins.

Method

Participants

Participants were 441 individual same-sex female twins (age range 16–25 years) who participated in the Twin Study of Hormones and Behavior across the Menstrual Cycle (HBMC) project (Klump et al., 2013, 2014) within the Michigan State University Twin Registiy (MSUTR; Bull & Klump, 2013; Klump & Burt, 2006). There are 265 families in the current study, of which there were 176 (66.42%) complete pairs (MZ = 105, DZ = 71) and 89 (33.58%) incomplete pairs (MZ = 40, DZ = 49). Study inclusion/exclusion criteria were: (a) menarche before the age of 15; (b) regular menstrual cycles every 22–32 days for the past 6 months; (c) no hormonal contraceptive use within the past 3 months; (d) no psychotropic or steroid medications within the past 4 weeks; (e) no pregnancy or lactation within the past 6 months; and (f) no history of genetic or medical conditions known to influence hormone functioning or appetite/weight. Although twins could have provided up to 45 days of affect data, data coverage beyond 30 days was low. All twins thus provided at least 17 consecutive days of data and a maximum of 30 consecutive days. The mean number of days of data contributed was 27.31 ($SD = 2.76$). Approximately three-quarters of the sample (75.28%) had one complete cycle within the 30 days. MZ twins contributed 0.72 days more than DZ twins ($t = 2.64$, $df = 439$, $p = .009$). Sample demographics are provided in Supplementary Table S1. Twins, on average, were in late adolescence ($M = 17.62$, $SD = 1.75$) and born to college-educated parents (modal category was bachelor's degree for both twins' mother and father) who combined make greater than $60,000 per year (59.80%). The ethnic composition of the sample was White (84.09%), African American (10.69%), Asian (0.24%), American Indian/Alaskan Native (0.48%), and mixed ethnicity (4.51%).

Procedures

All study measures and procedures were approved by the Michigan State University Institutional Review Board (IRB) and the University of Southern California IRB (UP-19-00623). Questionnaires were completed each evening (after 5:00 p.m.) using an online data system or pre-printed Scantron cards. Twins were instructed to submit their daily questionnaire responses online and to use Scantrons only in the event that they had issues with their computer or Internet connectivity. If twins used Scantrons, they were asked to bring them to the lab at their next in-person assessment so study staff could account for the daily data. There were 315 twins who completed their daily assessments on a Scantron card at least once during the study and 126 who completed all of their assessments online. Thus, 71.43% of the sample provided Positive and Negative Affect Schedule (PANAS) scores via Scantron at least once during the study. On days in which Scantrons were completed, negative affect scores were 0.81 units greater than on days in which ratings were submitted electronically ($SE = 0.14$, $t = 5.62$, $p < .001$). For positive affect, scores were 1.02 units greater on days in which Scantrons were used instead of electronic submission ($SE = 0.19$, $t = 5.29$, $p < .001$).

Measures

Positive and Negative Affect Schedule—The PANAS (Watson, Clark, & Tellegen, 1988) was used to assess daily levels of positive and negative affect. The positive affect scale includes 10 items that assess a range of positive emotions (e.g., interested, excited, alert, inspired), while the negative affect scale consists of 10 items that assess a range of negative emotions (e.g., distress, nervousness, irritability, fear). The degree to which each emotion was experienced on each day of data collection was rated on a 5-point scale, ranging from 1 = *very slightly/not at all* to 5 = *extremely*. The positive and negative affect scales have exhibited excellent internal consistency. McDonald's omega ranged from .85 to .90 for the 10 positive affect items over the 30 days ($M = .89$) and .80 to .85 ($M = .83$) for the 10 negative affect items. For each day, the sums of the 10 positive and 10 negative affect items were used.

Time scale—Selecting a time metric common to all twins was restricted to aligning twins by a common day in the menstrual cycle. Twins' PANAS scores were sorted so that all twins' scores on Day 1 were the first day of the follicular phase (the day following the final day of bleeding in twins' previous menstrual cycle). Menstrual cycle phase was coded using daily reports and daily hormone values from saliva samples. The first day of bleeding served as the graph anchor, and subsequent days were coded based on this anchor, as well as the overall length of each cycle (Klump et al., 2015).

Data analysis—We first estimated univariate MZ and DZ twin correlations of positive and negative affect and plotted heritability and environmental estimates based on these twin correlations across the 30 days. Conventional univariate twin modeling assumptions were made in estimating heritability and environmental estimates (Neale & Cardon, 1992). Additive genetic (A) components are the cumulative additive effect of genotype and are estimated by virtue of the fact that MZ twins share 100% of their genes and DZ twins share 50% of their genes, on average. Shared environmental (C) components are the

cumulative effect of any environment that makes twins reared in the same family more similar to one another and estimated under the assumption that shared environmental influences affect twins similarly regardless of genetic relatedness. Nonshared environmental (E) components are any environmental factor that makes twins different from one another, including measurement error. Conventional twin models make further assumptions that A, C, and E components are uncorrelated with one another, that they do not interact with one another, and that genetic relatedness is the product of parents' random mating strategies. To illustrate how the heritability of positive and negative affect fluctuates across days, we estimated the proportion of variability attributable to heritability (h^2), shared environments (c^2), and nonshared environments (e^2) for each of the 30 days.

Longitudinal structural models were then fit to the data using a multilevel structural equation modeling approach, which is equivalent to traditional structural equation modeling approaches (ML-SEM; McArdle & Prescott, 2005). We structured the data so that days were wide formatted while twin pairs remained in long format and nested within each family. We estimated a genetic simplex model (baseline model) and compared this model to a P→E model (research model). In the *genetic simplex model*, twins' affect scores on each day were decomposed into random variance components (subscript 1 for Twin 1 and subscript 2 for Twin 2):

$$P_{ft,1} = b_{0t} + w_{ab}A^b_{ft} + E^b_{ft} + w_{aw}A^w_{ft,1} + E^w_{ft,1}, \text{ and } P_{ft,2} = b_{0t} + w_{ab}A^b_{ft} + E^b_{ft} + w_{aw}A^w_{ft,2} + E^w_{ft,2}.$$

The phenotypic scores, P, of Twin 1 and Twin 2 in family, f, at day, t, are decomposed into between-family and within-family genetic and environmental factors. All biometric variables (A and E), have a mean of zero so that the expectation of the phenotypic mean $E[P]_{ft,i}$ equals the intercept, b_{0t}. At the between-family level, phenotypic scores are decomposed into a genetic factor (A^b) and an environmental factor (E^b) shared by both twins. At the within-family level, twins' scores are decomposed into a genetic factor (A^w) and an environmental factor (E^w) unique to twins raised in the same family. The total genetic effect, A, for each twin is equal to:

$$A_{ft,1} = w_{ab}A^b_{ft} + w_{aw}A^w_{ft,1}, \text{ and } A_{ft,2} = w_{ab}A^b_{ft} + w_{aw}A^w_{ft,2}.$$

The variances of the A^b and A^w factors are constrained to be equal. The weights, w, are fixed values that indicate the proportion of genetic information shared by each twin pair. In MZ pairs, w_{ab} equals 1 and w_{aw} equals 0 to satisfy the assumption that identical twins share all of their genes. In DZ pairs, w_{ab} equals 0.5 and w_{aw}, equals 0.5 to satisfy the assumption that fraternal pairs of twins share only half of their segregating genes, on average; the other half varies between them. The weights of the genetic factors between- and within-families were scaled so that $w^2_{ab} + w^2_{aw} = 1$, with the expectation that A^b and A^w are uncorrelated, $E\left[A^b_{ft}, A^w_{ft,i}\right] = 0$.

In order to present the P→E model, we begin by presenting the conventional genetic simplex model (Eaves et al., 1986). Between- and within-family genetic and environmental factors are correlated via first-order autoregressions ($t > 1$) as follows:

$$A^b_{ft} = a_{ARt,t-1}A^b_{ft-1} + uA^b_{ft}, \text{ and } E^b_{ft} = c_{ARt,t-1}E^b_{ft-1} + uE^b_{ft}, \text{ and } A^w_{ft,i} = a_{ARt,t-1}A^w_{ft-1,i} + uA^w_{ft,i}, \text{ and } E^w_{ft,i} = e_{ARt,t-1}E^w_{ft-1,i} + uE^w_{ft,i}.$$

The autoregressive coefficients ($a_{ARt,t-1}$, $c_{ARt,t-1}$ and $e_{ARt,t-1}$) of the genetic and environmental variables and their corresponding disturbances (uA^b_{ft}, uE^b_{ft}, $uA^w_{ft,i}$ and $uE^w_{ft,i}$) were freely estimated. The first-order autoregressive processes between the genetic and environmental factors meet the independence assumption in conventional twin models. Graphically, this model is identical to Figure 1 *without* the bolded red pathways projecting from the phenotypic scores to the nonshared environmental variables. Pre-analysis demonstrated that the longitudinal genetic and shared environmental structures could each be reduced to a single common factor; in all of our models the $a_{ARt,t-1}$ and $c_{ARt,t-1}$ were not estimated. Reduced models that included single common genetic and shared environmental latent variables fit the data better than the full genetic simplex model (positive affect: ΔLL = 85.10, $df = 117$, $p = .988$; negative affect: ΔLL = 85.10, $df = 117$, $p > .999$).

In the P→E model, the autoregressive correlations of the E^w components are re-specified as regressions of $E^w_{ft,i}$ at Day $t > 1$ on the phenotype at Day t—1:

$$E^w_{ft,i} = b_{PEt,t-1}P^w_{t-1} + uE^w_{ft,i}.$$

The nonshared environmental variables, $E^w_{ft,i}$, at Day t>1, are regressed on the phenotype at Day t—1. The autoregressive coefficients ($b_{PEt,t-1}$) and disturbances $\left(uE^w_{ft,i}\right)$ were freely estimated. By path tracing rules, in the DZ group the $b_{PEt,t-1}$ parameters necessarily transmit effects of the within-family genetic factor $\left(A^w_{ft-1,i} * b_{PEt,t-1}\right)$ and the nonshared environmental factor $\left(E^w_{ft-1,i} * b_{PEt,t-1}\right)$ at Day t—1 via the phenotype at Day t—1. As there is no within-family genetic variation, A^w variables are fixed to zero in the MZ group, thus making $b_{PEt,t-1}$ effects equal to the autoregressive effects of the nonshared environment, $e_{ARt,t-1}$ in the MZ group. The baseline model, therefore, assumes that $b_{PEt,t-1}$ effects are different in the MZ group than the DZ group, as the source variables for $b_{PEt,t-1}$ effects include within-family genetic differences in the DZ twins but not the MZ twins. In the DZ group, significant $b_{PEt,t-1}$ effects support the hypothesis that pair differences in affect on previous days systematically expose twins to unique environments on subsequent days, which causes unique environments to become correlated with genetic factors and unique environments on previous days. In the MZ group, significant $b_{PEt,t-1}$ effects suggest that unique environments at Day t correlate with unique environments at t—1.

The baseline model was compared to a restricted model that constrains all $b_{PEt,t-1}$ parameters to be equal between MZ and DZ twins. If the more restricted model fits equally as well as the baseline model, this means that $b_{PEt,t-1}$ parameters cannot be

distinguished from the $e_{ARt,t-1}$. Under this more parsimonious model, the null hypothesis is that environmental effects over time are equivalent to person–environment effects over time.

The P→E parameter, $b_{PEt,t-1}$, has three consequences in the DZ group. First, it necessarily induces accumulating within-family *r*GE over days. Second, $b_{PEt,t-1}$ parameters change the meaning of the nonshared environmental effects when $t > 1$ because the parameters give rise to the correlations between $E^{w}_{ft,i}$ and $A^{w}_{ft-1,i}$ (Dolan et al., 2014). Third, $b_{PEt,t-1}$ parameters increase the stability of nonshared environment over days as twins are matched to temporally similar environments.

Since b_{PE} parameters necessarily change the meaning of the nonshared environment at all days but the first day via indirect correlations between genetic and environmental components (Beam et al., 2015), we empirically test the longitudinal structure of the nonshared environment by statistically comparing model implied nonshared environmental variances and covariances in the DZ group against the MZ group.

Models were fit in M*plus* 8.2 (Muthén & Muthén, 1998–2017) using full information maximum likelihood (FIML) estimation with robust standard errors to handle missing data, violations of multivariate normality, and modest sample sizes. Missing data analysis consisted of comparing participants who did not provide affect scores for all 30 days to participants who did on the following measures: first day of affect scores provided, Day 15 of affect scores, final day of affect scores provided, age, body mass index, highest education achieved, parental income, and ethnicity. No significant group differences were observed, which lends support for retaining the assumption that the missing data mechanism was missing completely at random (Enders, 2010). Model comparisons were made with the Satorra-Bentler scaled chi-square difference test of nested models (S-Bχ2), which corrects the chi-square distributed test statistic in cases of multivariate normality assumption violations (Satorra & Bentler, 2001). Model fit also was evaluated using the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) (Burnham & Anderson, 2004). We also employed the sample size adjusted Bayesian information criterion (SSABIC), which adjusts the sample size penalty of the BIC to provide better model fit performance than the BIC when sample sizes are small or there are a large number of parameters (Enders & Tofighi, 2007; Sclove, 1987). Models with lower AIC, BIC, and SSABIC values indicate better model fit, because these indices take into account the tradeoff between model parsimony and model complexity.

Results

The means and standard deviations of positive and negative affect across the 30 days were stable, with greater variance in positive affect (range: 21.95–24.75) than negative affect (range: 14.55–15.56).

Figure 2 presents the twin correlations for positive affect and negative affect on each day with locally estimated scatterplot smoothing (LOESS) lines overlaid to illustrate trends in twin similarity over days. Beginning at Day 1, MZ and DZ twin correlations for positive affect (top panel) gradually diverge until about Day 10, remain constant from Day 11

through Day 18, diverge again from Day 18 through about Day 28, and then converge the final two days. For negative affect (bottom panel), twin correlations generally converge from Day 1 through Day 11, then diverge until about Day 28, after which the correlations tend to converge with DZ twins more similar than MZ twins in the final days. Within-pair differences in person–environment match, the basis for generating within-family *r*GE, are expected to account for divergence of MZ and DZ twin correlations.

The above twin correlations were used to estimate the daily heritability and environment estimates (Figure 3). Both MZ and DZ correlations demonstrated variability in twin similarity in positive affect scores over the 30 days. Heritability estimates (red lines) tended to increase slightly until about Day 14, were variable from Day 15 through Day 19, increased dramatically from Day 19 through Day 25, and then declined. Shared environmental estimates were generally around zero (or negative), although on days where heritability declined, shared environmental estimates tended to increase above zero (e.g., Days 6, 14, 29, and 30). Nonshared environmental estimates over days were stable.

Twin correlations for negative affect also tended to vary across days, suggesting variability in genetic and environmental estimates. Heritability estimates generally declined from Day 1 through Day 12, increased through Day 25, and then slightly declined. Shared environmental estimates tended to increase across days where heritability declined and decreased on days when heritability increased. Nonshared environmental estimates slightly increased for the first 15 days and slightly declined for the last 15 days. As increases in heritability suggest that DZ twin similarity diverges from MZ twin similarity, we next tested whether *r*GE accounts for daily fluctuations in heritability, particularly across days where heritability of positive and negative affect increases.

Person→environment model results

Model fitting results (Table 1) indicate that the P→E model provided better fit to the data than a model that equates P→E and nonshared environmental parameters for both positive affect (likelihood ratio test (LRT) = 86.61, $df = 29$, $p < .001$) and negative affect (LRT = 53.49, $df = 29$, $p < .005$). Further, AIC and SSABIC values were lower for models that distinguish P→E parameters than models that do not, further indicating better fit to the data. Person–environment matching parameters (b_{PE}), thus, significantly improved model fit for both positive affect and negative affect. Results did not change when ethnicity and age were included in the models.

Figure 4 presents the line plots of model estimated within-family *r*GE for positive affect and negative affect from the P→E models. Within-family *r*GE was small and tended to fluctuate from Day 1 through Day 15. From Day 15 through Day 25, *r*GE increased but began to decline again from Day 26 through Day 30.

In order to evaluate whether *r*GE fluctuated systematically as a function of daily effects of affect scores on subsequent environmental exposure, we randomly ordered days within each twin and fit the P→E model. This analysis, known as surrogate data generation or “time scrambling” (Moulder, Boker, Ramseyer, & Tschacher, 2018), tests dependency between time series and is important for showing that the model-generated *r*GE depends on the

temporal ordering of P→E effects. Thus, the expectation is that *r*GE would be essentially zero across the 30 days when days were randomly ordered within twin. Within-family *r*GE did not vary significantly from zero when days were randomly ordered within twins (lines labeled “PA Random Days” and “NA Random Days”, respectively).

The P→E parameters affected heritability of daily affect scores attributed to genetic effects alone, that is, estimates excluding *r*GE generated via b_{PE} paths (Figure 5). Heritability estimates under the independence assumption (blue lines) for positive affect and for negative affect are compared to heritability estimates under accommodation of *r*GE (red lines). For both positive and negative affect, estimating P→E effects in the DZ group reduced heritability estimates by approximately 3% across days, on average, with reductions as great as 7–8%. Heritability estimates, thus, were attenuated when *r*GE was included in the model.

Finally, P→E effects that accommodate *r*GE also affected the longitudinal correlations between nonshared environmental components across the 30 days. Figure 6 presents the nonshared environmental correlations in the MZ group (left column) and the DZ group (right column) for positive affect (top row) and negative affect (bottom row). As a reminder, longitudinal correlations between nonshared environmental components in the MZ group do not include *r*GE (because there are no within-family genetic effects) whereas in the DZ group *r*GE is accommodated. In the MZ group, longitudinal nonshared environmental correlations persist up to 5 days, declining from a moderate correlation of .40–.50 between adjacent days (e.g., Day 1 and Day 2) to essentially zero 5 days later (e.g., between Day 1 and Day 6). Nonshared environmental correlations decayed more rapidly when accommodating *r*GE by about 1–1.5 days earlier, but are systematically correlated across days, although the correlations are small. Differences in longitudinal correlations estimated with and without accommodation of *r*GE are similar, suggesting that the interpretation of nonshared environmental components may have been affected minimally by allowing genetic and environmental components to correlate. Differences in lag-1 correlations were the largest (*r* range: .00−.20) while correlations among components with greater lags never exceeded .10.

Post-hoc power analysis

Given the small total sample size (N_{Families} = 265), we estimated power to detect significant P→E parameters, $b_{PEt,t-1}$, across the 30 days. Power was estimated using the Markov Chain Monte Carlo feature in in M*plus* 8.2 (Muthén & Muthén, 1998–2017) using the final values from the P→E models for positive and negative affect. Two thousand replications were specified for each analysis. For positive affect, power to detect significant $b_{PEt,t-1}$ parameters ranged from .39–1.00 (*M* = .86, *SD* = .16; *Mdn* = .94). For negative affect, power to detect significant $b_{PEt,t-1}$ parameters ranged from .26–1.00 (*M* = .70, *SD* = .30, *Mdn* = .90).

Discussion

Plomin et al. (1977) recognized that “environmental and genetic threads in the fabric of behavior are so tightly interwoven that they are indistinguishable” (p. 309). The twin data presented here and elsewhere (Beam et al., 2015; Dolan et al., 2014) suggest that genetic and environmental influences covary across time (i.e., days and years) to support differentiation

in psychological traits and abilities over time. As matching between individuals and their environments ebbs and flows, so does *r*GE. These results, thus, address one process through which environments might come to be correlated with outcomes and genotype (Plomin, 1986).

Modeling approaches for accommodating and testing *r*GE (Beam & Turkheimer, 2013; de Kort et al., 2012; Dolan, Huijskens, Minică, Neale, & Boomsma, 2019; Moscati, Verhulst, McKee, Silberg, & Eaves, 2018) are finally catching up to discussions over how genetic and environmental influences reciprocally influence one another (Anastasi, 1958; Briley et al., 2019; Eaves et al., 1977; Wachs, 1983). The REM approach taken in the current study places the individual at the center of change processes that drive observed phenotypic differences in positive and negative affect. Individuals are not passive recipients of their genotype and environments, as is implied in conventional twin models through specification of unidirectional pathways from genetic and environmental variance components to phenotypic outcomes. In this way, phenotypic stability not only depends on the degree of genetic and environmental relatedness but also on the consistency of similarity and differences of individuals' environmental systems (Lickliter & Harshaw, 2010; West & King, 1987). Siblings, for example, could increasingly differ if they are exposed to different learning activities and environments (e.g., one child is read to daily while the other is not) yet converge if their environments are highly similar (e.g., whatever one child gets, so does the other). By including person–environment effects in the current model that quantify the strength of the relation between twins' affective outcomes and unique environments, we have shown real time selection and evocation of environments that contribute to daily differences in twins' affect.

Person–environment effects that induce within-family *r*GE have implications for heritability estimates and longitudinal correlations of nonshared environments. In the present study, we found lower heritability estimates when allowing indirect relations between twins' unique genotype and their unique environments, compared to models in which independence between genotype and environment was assumed. Adaptation to one's unique environmental context, thus, might account for why affect scores appeared to be highly heritable at first glance. Lower heritability estimates when *r*GE is modeled are a well-known statistical consequence due to *r*GE inflating genetic variance (Briley et al., 2019; Purcell, 2002) and possibly heritability, at least under the assumption of constant genetic variance across a range of environments.

We found support for the hypothesis that genetic influences on affect come to be correlated with environments because of the correlation between phenotypes and subsequent environmental influences contributing to their phenotypes (Gottlieb, 2003). Female twins in the sample may not have been selecting and reacting to environments randomly; rather, the data are consistent with our supposition that individuals select some environments and react to others based on the suitability of those environments. Further, given twins' relatively stable affect scores across the 30 days, one possibility is that person–environment matching processes rise and fall over time with the ultimate purpose of stabilizing positive and negative affect responses over time.

Nonshared environments are expected to correlate across days, as environmental systems are not randomly distributed over time, although they tend to decay quickly over time (Burt, Khlar, & Klump, 2015). The content of environmental systems is worthy of comment. As nonshared environmental components in twin models comprise any nongenetic feature that contributes to within pair differences, including measurement error, identification of specific environmental factors that correlate with twins' affect scores is beyond the scope of this study. Environmental systems probably comprise twins' different interpretations of their shared relationships (e.g., parents) or neighborhoods, their unique peer networks, hormone levels (both stress and sex hormones), and molecular environments such as differences in DNA methylation and RNA transcription (Gottlieb, 2003). Twins who prefer quiet solitude compared to their co-twins, for example, probably avoid loud and crowded venues no matter the situation, but especially if they are perceived as engendering stress that might increase negative affect or lower positive affect. As Scarr (1992) put it, environments are "largely the construction of individual family members in the ways they evoke responses from others, actively select or ignore opportunities, and construct their own experiences" (p.14). Construction of and reaction to experiences, broadly construed, should correlate with genotype, as genotype is responsive to environmental experiences (Kendler & Baker, 2007). Modeled *r*GE, thus, showed that twins' total unique environments are temporally linked for longer than expected under models that assume genetic and environmental covariation.

Of note, *r*GE estimates fluctuated systematically over time, particularly between Days 15 and 30 – the latter "half" of the menstrual cycle. As the current study is observational, we cannot rule out these and other third variable confounds that account for P→E effects (e.g., coping skills, distress, physical illness/disability). Although we could not identify specific biological or environmental factors in twins' lives that determine increases in *r*GE underlying positive and negative affect, we can only speculate about the reasons why. Stress vulnerability to daily hassles, possibly because of hormonal factors, social environmental factors, or both, might be greater in latter days of the cycle. Previous findings suggest that the luteal phase – and particularly the late luteal period – may represent a period of intensified daily hassles compared to the follicular phase (Kiesner, Mendle, Eisenlohr-Moul, & Pastore, 2016). Women may be more prone to select environments for stress management, as environmental factors such as perceived stress and social support might exacerbate cycle-related changes in affect (Romans et al., 2012). Accordingly, environmental context has been identified as an important contributor to daily mood ratings (Stone, Marco, Cruise, Cox, & Neale, 1996). A successful match between persons and environments, thus, might be especially relevant during high-stress vulnerability phases relative to low-stress vulnerability phases.

Environmental components in the within-family *r*GE estimates likely consist of the correlation between genotype and hormones. Estrogen and progesterone trigger changes in gene expression (Cole, 2009; Östlund, Keller, & Hurd, 2003), and genetic factors account for individual physiological and behavioral responses to estrogen and progesterone levels (Klump et al., 2007; Wall et al., 2014). Under this explanation, any influence ovarian hormones have on genotype would not make much of a difference for hypothesized P→E matching. Hormones, rather, are a source variance that accounts for genetic variability that ultimately correlates with nonshared environmental factors downstream. Within-pair

differences in estrogen and progesterone levels during the luteal phase, however, might moderate (or mediate) within-pair genetic effects on differences in affect, as was found for binge eating in the same sample (Klump, Fowler, Mayhall, Sisk, Culbert, & Burt, 2018). Significant P→E effects, thus, may reflect Gene×Environment interaction (GxE) rather than *r*GE. P→E models cannot rule out GxE processes as the P→E generating mechanism or whether the GxE processes are exogenous or endogenous to individuals. In reality, we expect *r*GE and GxE processes to contribute to observed P→E effects. In general, indirect and bidirectional associations among molecular factors, like ovarian and stress hormones, genotype, and environment are more likely than direct genetic effects on affect, similar to most, if not all, human complex traits (Anastasi, 1958; Gottlieb, 2003).

Future research implementing P→E parameters, thus, would benefit from specifying measured biological (e.g., mRNA and hormones) and environmental (e.g., peer relationships) factors, in addition to nonspecific environmental factors (Wachs, 1983). Future longitudinal twin modeling would also benefit from incorporating GxE in the presence of *r*GE (Johnson, 2007). P→E models that include intermediate variables in the causal chain and GxE processes need to be developed.

Although accommodating *r*GE in longitudinal behavioral genetic models has the advantage of explicating how environments come to be correlated with psychological traits and heredity, P→E parameters change the meaning of nonshared environmental components (Dolan et al., 2014). Our empirical test compared nonshared environmental components with and without *r*GE and suggested that including P→E may do so only minimally. While the theoretical meaning of the nonshared environment does change by virtue of genetic components indirectly predicting environmental components, the model-estimated non-shared environment correlations did not differ appreciably, suggesting overall similarity in the interpretation of the nonshared environment. While twin studies do not clarify specific environmental factors that contribute to any phenotype (Wachs, 1983), the implication that differences in environmental exposure matter for maintaining stability of affect seem to be preserved when genetic and environmental components are allowed to correlate.

Accommodating *r*GE in longitudinal twin models explicitly demonstrates how the mutual exchange between people and their environments might cause highly heritable traits to increase over time, as originally proposed by Dickens and Flynn (2001). Meta-analyses of longitudinal findings have drawn on Dickens and Flynn's version of the REM to understand age trends in genetic and environmental variance components, but only indirectly (Briley & Tucker-Drob, 2013, 2014; Tucker-Drob & Briley, 2014). These studies must assume the same independence assumption and so postulate that twin correlations decrease with age in DZ twins compared to MZ twins (see also McCartney, Harris, & Bernieri, 1990), but never address the developmental systems that explain *how* DZ twins diverge in similarity over time.

The temporal dynamics of phenotype–environment matching might differ across traits and abilities. For traits that tend to be highly stable and evolve slowly over time, like personality (Roberts & Mroczek, 2008), within-family *r*GE is expected to rise precipitously and to level off when persons are more or less canalized into their adult environments.

Research designs probably can include long intervals of time between measurements for such phenotypes and still detect within-family *r*GE, as environments that support these traits are unlikely to change over short intervals of time and are expected to be highly correlated and decline slowly (e.g., jobs tend to consist of similar environmental demands over long periods). For phenotypes that fluctuate daily, such as affect, perceived stress, and psychiatric symptoms (e.g., depressive symptomatology), within-family *r*GE is expected to increase and decrease over relatively short intervals of time because of state dependence. Further, nonshared environmental correlations would be expected to persist, but decline quickly with the passage of time. Individuals prone to depression, for example, are not continuously depressed nor do individuals prone to seeing their world as stressful always feel stressed. Minor arguments with a friend that influence depression and stress symptoms one day may have short-term effects on environments the next day or the following day (e.g., distancing oneself to calm down), but generally do not persist over many days (e.g., interpersonal differences are mended after a few days).

Consistent with the above expectations, etiological influences on positive and negative affect, including *r*GE, fluctuated across days rather than increased linearly. Detection of within-family *r*GE for state-dependent phenotypes, thus, requires shorter divisions of time that can capture day-to-day variability in *r*GE. This is the first study to test *r*GE of a phenotype over days, and demonstrates that the *r*GE underlying state-dependent phenotypes are observed when lags between measurements are short. In one previous study with depressive symptom measurements separated by years, person–environment effects that generate within-family *r*GE were not found (Beam et al., 2016), possibly because the intervals of time between measurements (approximately three years) were too long to capture meaningful person–environment correlations underlying the temporal dynamics of depressive symptoms. By comparison, in studies of phenotypes that are relatively stable over years, like cognitive ability (Beam et al., 2015; de Kort et al., 2014) and personality (Beam & Sharp, 2020), within-family *r*GE was observed and tended to increase monotonically and stabilize over time.

Limitations & future directions

The most notable limitation in the current study is that processes other than *r*GE (e.g., GxE interaction) might account for stability of pair differences in affect over the 30 days of the menstrual cycle studied here. Along this same theme, a second limitation is that other biological processes that factor into pair differences remain unaccounted for in the current study. Molecular environments, like polygenic scores (Dolan et al., 2019), RNA transcription and protein transcription, responsible for genetic expression have a great deal to do with phenotypic expression (Cole, 2009; Gottlieb, 2003). Intermediary pathways that lie between genotype, behavior, and environments may help to clarify how heritability changes over time.

A third limitation of the current study is the assumption that all twins experienced functionally equivalent environments across the 30 days, such that the roles of specific environments cannot be made (Wachs, 1983). In other words, it remains unclear whether one twin experienced a more extreme environment (e.g., trauma exposure) compared to

her co-twin. Likewise, nonshared environmental variance and measurement error were not differentiated, so we cannot conclude definitively that genetic variance correlates with nonshared environmental sources of variance alone.

A fourth limitation is that we did not account for the possibility that genetic variance underlying affect at Day *t* might also correlate with shared environmental variance underlying affect at Day *t*+1. While it is possible to model between-family *r*GE using a similar approach as the one taken in the current study, such modeling is useful for testing how between-family *r*GE causes both twins in a family to diverge from both twins in another family. As our focus was on whether within-family *r*GE partially explains how individuals diverge, modeling between-family *r*GE fell outside our study goals.

A fifth limitation is that the P→E model assumes that the autocorrelation structure between measurement occasions are stationary and that the process under scrutiny is noncyclic. The time metric used in the current study, the menstrual cycle, however, is nonstationary and cyclic. The P→E parameterization, thus, may oversimplify the person–environment matching process hypothesized to occur over time in these data. As there are no other daily twin studies with as many repeated measurements as in the MSUTR, this study offers a rare glimpse into person–environment processes that account for differences in daily affect scores.

A final limitation concerns the study design. First, the sample size was relatively small with only 265 families. Yet, the repeated measurement design increases measurement precision, which in turn reduces the necessary sample size to achieve a given level of power (Allison et al., 1998; Evans, 2002). This was reflected in the adequate power to detect P→E effects, on average, observed in our post-hoc power analysis. Second, we observed significant within-person differences in positive and negative affect scores between days in which affect scores were supplied online versus supplied on Scantron cards, with slightly higher scores reported on days using Scantron cards. This limitation, however, is offset by the advantage of recording twins' scores rather than not on days when twins did not have online access.

There are a number of directions for future research that would benefit the field of developmental behavioral genetics, both of which are oriented toward understanding *how* heritability changes with time and become correlated with nonshared environments. First, daily twin studies with time scales that are both stationary and noncyclic should be a priority, as this study design would satisfy the P→E model assumptions. Second, the inclusion of measured genotypic differences, like polygenic scores, would permit tests of whether nonshared environmental differences correlate with differences in measured causal variants. Work in this area is underway (Dolan et al., 2019). A third direction involves cohort sequential designs in longitudinal twin studies. This approach would help clarify when P→E effects are greatest across different developmental periods. Certain periods of development might be more critical for person–environment matching than others (e.g., looking for a first job vs. a lateral career move in midlife). The fourth direction we hope to see in future studies is the integration of experimental designs with behavioral genetic approaches, like the one recently taken by Burt, Plaisance, and Hambrick (2019). Randomly assigning twins within pairs to different environmental exposures may offer insight into the ways in

which twins' construction and reaction to initial random environmental exposure influences developmental trajectories and *r*GE. Random assignment to different environments may induce twins to differentiate within brief windows of time because of within-family *r*GE processes, provided that the trait, like affect, can be manipulated over short periods of time. Although P→E effect sizes might be small (see also Dolan et al., 2014), they can have meaningful consequences for developmental outcomes when repeated consistently over time (Funder & Ozer, 2019).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknoweledgement.

We thank Eric Turkheimer for providing helpful comments on an early draft of this manuscript.

Funding Statement. This work was supported by the National Institute of Mental Health (R01 MH082054) and the National Institute on Aging (T32 AG000037-37).

References

Allison DB, Thiel B, St. Jean P, Elston RC, Infante MC, & Schork NJ (1998). Multiple phenotype modeling in gene-mapping studies of quantitative traits: Power advantages. American Journal of Human Genetics, 63, 1190–1201. doi: 10.1086/302038 [PubMed: 9758596]

Anastasi A (1958). Heredity, environment, and the question "how?". Psychological Review, 65, 197–208. [PubMed: 13579088]

Bäckström T, Haage D, Löfgren M, Johansson IM, Strömberg J, Nyberg S, … Bengtsson SK (2011). Paradoxical effects of GABA-A modulators may explain sex steroid induced negative mood symptoms in some persons. Neuroscience, 191, 46–54. doi: 10.1016/j.neuroscience.2011.03.061 [PubMed: 21600269]

Bartels M, Rietveld MJH, Van Baal GCM, & Boomsma DI (2002). Genetic and environmental influences on the development of intelligence. Behavior Genetics, 32, 237–249. [PubMed: 12211623]

Bartels M, van den Oord EJCG, Hudziak JJ, Rietveld MJH, van Beijsterveldt CEM, & Boomsma DI (2004). Genetic and environmental mechanisms underlying stability and change in problem behaviors at ages 3, 7, 10, and 12. Developmental Psychology, 40, 852–867. doi: 10.1037/0012-1649.40.5.852 [PubMed: 15355171]

Beam CR, Emery RE, Reynolds CA, Gatz M, Turkheimer E, & Pedersen NL (2016). Widowhood and the stability of late life depressive symptomatology in the Swedish adoption twin study of aging. Behavior Genetics, 46, 100–113. doi: 10.1007/s10519-015-9733-7 [PubMed: 26303346]

Beam CR, & Sharp ES (2020). Social dynamics of personality development in late adulthood. Manuscript in preparation.

Beam CR, & Turkheimer E (2013). Phenotype–environment correlations in longitudinal twin models. Development and Psychopathology, 25, 7–16. doi: 10.1017/S0954579412000867 [PubMed: 23398749]

Beam CR, & Turkheimer E (2017). Gene-environment correlation as a source of stability and diversity in development. In Tolan PH, & Leventhal B (Eds.), Gene-environment transactions in developmental psychopathology: Role in intervention research (pp. 111–130). New York: Springer.

Beam CR, Turkheimer E, Dickens WT, & Davis DW (2015). Twin differentiation of cognitive ability through phenotype to environment transmission: The Louisville Twin Study. Behavior Genetics, 45, 622–634. doi: 10.1007/s10519-015-9756-0 [PubMed: 26468113]

Bell RQ (1968). A reinterpretation of the direction of effects in studies of socialization. Psychological Review, 75(2), 81–95. [PubMed: 4870552]

Boker SM, & McArdle JJ (2014). Path analysis and path diagrams. Retrieved from https://onlinelibrary.wiley.com/doi/10.1002/9781118445112.stat06517

Boomsma DI, & Molenaar PCM (1987). The genetic analysis of repeated measures. I. simplex models. Behavior Genetics, 17, 111–123. [PubMed: 3606535]

Briley DA, Livengood J, Derringer J, Tucker-Drob EM, Fraley RC, & Roberts BW (2019). Interpreting behavior genetic models: Seven developmental processes to understand. Behavior Genetics, 49(2), 196–210. doi: 10.1007/s10519-018-9939-6 [PubMed: 30467668]

Briley DA, & Tucker-Drob EM (2013). Explaining the increasing heritability of cognitive ability across development: A meta-analysis of longitudinal twin and adoption studies. Psychological Science, 24, 1704–1713. doi: 10.1177/0956797613478618 [PubMed: 23818655]

Briley DA, & Tucker-Drob EM (2014). Genetic and environmental continuity in personality development: A meta-analysis. Psychological Bulletin, 140, 1303–1331. [PubMed: 24956122]

Bronfenbrenner U, & Ceci SJ (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. Psychological Review, 101, 568–586. doi: 10.1037/0033-295X.101.4.568 [PubMed: 7984707]

Burnham KP, & Anderson DR (2004). Multimodel inference: Understanding AIC and BIC in model selection. Sociological Methods & Research, 33, 261–304. doi: 10.1177/0049124104268644

Burt SA, Klahr AM, & Klump KL (2015). Do non-shared environmental influences persist over time? An examination of days and minutes. Behavior Genetics, 45(1), 24–34. [PubMed: 25262214]

Burt SA, & Klump KL (2013). The Michigan State University Twin Registry (MSUTR): An update. Twin Research and Human Genetics, 16, 344–350. doi: 10.1017/thg.2012.87 [PubMed: 23101567]

Burt SA, Plaisance KS, & Hambrick DZ (2019). Understanding “what could be”: A call for “experimental behavioral genetics”. Behavior Genetics, 49, 235–243. doi: 10.1007/s10519-018-9918-y [PubMed: 30105514]

Cole SW (2009). Social regulation of human gene expression. Current Directions in Psychological Science, 18, 132–137. [PubMed: 21243077]

de Kort JM, Dolan CV, & Boomsma DI (2012). Accommodation of genotype-environment covariance in a longitudinal twin design. Netherlands Journal of Psychology, 67, 81–90.

de Kort JM, Dolan CV, Kan K-J, van Beijsterveldt CEM, Bartels M, & Boomsma DI (2014). Can GE-covariance originating in phenotype to environment transmission account for the Flynn effect? Journal of Intelligence, 2, 82–105. doi: 10.3390/jintelligence2030082.

Dickens WT, & Flynn JR (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. Psychological Review, 108, 346–369. [PubMed: 11381833]

Dickens WT, Turkheimer ET, & Beam CR (2011). The social dynamics of the expression of genes for cognitive ability. In Kendler KS, Jaffee S, & Romer D (Eds.), The dynamic genome and mental health: The role of genes and environments in youth development (pp. 103–127). New York: Oxford University Press.

Dolan CV, De Kort JM, Van Beijsterveldt TCEM, Bartels M, & Boomsma DI (2014). GE covariance through phenotype to environment transmission: An assessment in longitudinal twin data and application to childhood anxiety. Behavior Genetics, 44, 240–253. doi: 10.1007/s10519-014-9659-5 [PubMed: 24789102]

Dolan CV, Huijskens RCA, Minică CC, Neale MC, & Boomsma DI (2019). Incorporating polygenic scores in the twin model to estimate genotype-environment covariance: exploration of statistical power. bioRxiv, 1–32. doi: 10.1101/702738

Eaves LJ, Krystyna L, Martin NG, & Jinks JL (1977). A progressive approach to non-additivity and genotype-environmental covariance in the analysis of human differences. British Journal of Mathematical and Statistical Psychology, 30, 1–42.

Eaves LJ, & Long J, & Heath AC. (1986). A theory of developmental change in quantitative phenotypes applied to cognitive development. Behavior Genetics, 16, 143–162. [PubMed: 3707482]

Enders CK (2010). Applied missing data analysis. New York, NY: The Guilford Press.

Enders CK, & Tofighi D (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. Psychological Methods, 12(2), 121–138. [PubMed: 17563168]

Evans DM (2002). The power of multivariate quantitative-trait loci linkage analysis is influenced by the correlation between variables. American Journal of Human Genetics, 70, 1599–1602. doi: 10.1086/340850 [PubMed: 11992269]

Farage MA, Neill S, & MacLean AB (2009). Physiological changes associated with the menstrual cycle a review. Obstetrical and Gynecological Survey, 64, 58–72. doi: 10.1097/OGX.0b013e3181932a37 [PubMed: 19099613]

Fischbein S (1978). Heredity-environment interaction in the development of twins. International Journal of Behavioral Development, 1, 313–322. doi: 10.1177/016502547800100402

Fuller JL, & Thompson WR (1960). Behavior genetics. New York: John Wiley & Sons, Inc.

Funder DC, & Ozer DJ (2019). Evaluating effect size in psychological research: Sense and nonsense. Advances in Methods and Practices in Psychological Science, 2(2), 156–168.

Gottlieb G (2003). On making behavioral genetics truly developmental. Human Development, 46, 337–355. doi: 10.1159/000073306

Halbreich U, Freeman EW, Rapkin AJ, Cohen LS, Grubb GS, Bergeron R, … Constantine GD (2012). Continuous oral levonorgestrel/ethinyl estradiol for treating premenstrual dysphoric disorder. Contraception, 85, 19–27. doi: 10.1016/j.contraception.2011.05.008 [PubMed: 22067793]

Johnson W (2007). Genetic and environmental influences on behavior: Capturing all the interplay. Psychological Review, 114, 423–440. doi: 10.1037/0033-295X.114.2.423 [PubMed: 17500633]

Kendler KS, & Baker JH (2007). Genetic influences on measures of the environment: A systematic review. Psychological Medicine, 37, 615–626. doi: 10.1017/S0033291706009524 [PubMed: 17176502]

Kiesner J, Mendle J, Eisenlohr-Moul TA, & Pastore M (2016). Cyclical symptom change across the menstrual cycle: Attributional, affective, and physical symptoms. Clinical Psychological Science, 4(5), 882–894.

Klump KL, & Burt SA (2006). The Michigan State University Twin Registry (MSUTR): Genetic, environmental and neurobiological influences on behavior across development. Twin Research and Human Genetics, 9, 971–977. doi: 10.1375/183242706779462868 [PubMed: 17254439]

Klump KL, Burt SA, Mcgue M, & Iacono WG (2007). Changes in genetic and environmental influences on disordered eating across adolescence. Archives of General Psychiatry, 64, 1409–1415. [PubMed: 18056549]

Klump KL, Fowler N, Mayhall L, Sisk CL, Culbert KM, & Burt SA (2018). Estrogen moderates genetic influences on binge eating during puberty: Disruption of normative processes? Journal of Abnormal Psychology, 127(5), 458–470. [PubMed: 29927265]

Klump KL, Hildebrandt BA, O'Connor SM, Keel PK, Neale M, Sisk CL, … Burt SA (2015). Changes in genetic risk for emotional eating across the menstrual cycle: A longitudinal study. Psychological Medicine, 45(15), 3227–3237. [PubMed: 26174083]

Klump KL, Keel PK, Racine SE, Burt AA, Neale M, Sisk CL, … Hu JY (2013). The interactive effects of estrogen and progesterone on changes in emotional eating across the menstrual cycle. Journal of Abnormal Psychology, 122, 131–137. doi: 10.1037/a0029524 [PubMed: 22889242]

Klump KL, Racine SE, Hildebrandt B, Burt SA, Neale M, Sisk CL, … Keel PK (2014). Influences of ovarian hormones on dysregulated eating: A comparison of associations in women with versus women without binge episodes. Clinical Psychological Science, 2(5), 545–559. [PubMed: 25343062]

Lickliter R, & Harshaw C (2010). Canalization and malleability reconsidered: The developmental basis of phenotypic stability and variability. In Hood KE, Halpern CT, Greenberg G & Lerner RM (Eds.), Handbook of developmental science, behavior, and genetics (pp. 491–525). West Sussex: Wiley-Blackwell.

McArdle JJ, & Prescott CA (2005). Mixed-effects variance components models for biometric family analyses. Behavior Genetics, 35, 631–652. doi: 10.1007/s10519-005-2868-1. [PubMed: 16184490]

McCartney K, Harris MJ, & Bernieri F (1990). Growing up and growing apart: A developmental meta-analysis of twin studies. Psychological Bulletin, 107, 226–237. [PubMed: 2138795]

McGue M, Osler M, & Christensen K (2010). Causal inference and observational research: The utility of twins. Perspectives on Psychological Science, 5, 546–556. doi: 10.1177/1745691610383511. [PubMed: 21593989]

Montag C, Hahn E, Reuter M, Spinath FM, Davis K, & Panksepp J (2016). The role of nature and nurture for individual differences in primary emotional systems: Evidence from a twin study. PLoS ONE, 11, 1–15. doi: 10.1371/journal.pone.0151405.

Moscati A, Verhulst B, McKee K, Silberg J, & Eaves L (2018). Cross-lagged analysis of interplay between differential traits in sibling pairs: Validation and application to parenting behavior and ADHD symptomatology. Behavior Genetics, 48, 22–33. doi: 10.1007/s10519-017-9882-y. [PubMed: 29150722]

Moulder RG, Boker SM, Ramseyer F, & Tschacher W (2018). Determining synchrony between behavioral time series: An application of surrogate data generation for establishing falsifiable null-hypotheses. Psychological Methods, 23, 757–773. doi: 10.1037/met0000172 [PubMed: 29595296]

Muthén LK, & Muthén BO (1998–2017). Mplus user's guide. (8th ed.). Los Angeles, CA: Muthén & Muthén.

Neale MC, & Cardon LR (1992). Methodology for genetic studies of twins and families. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Neale MC, & McArdle JJ (2000). Structured latent growth curves for twin data. Twin Research and Human Genetics, 3(3), 165–177.

Nivard MG, Dolan CV, Kendler KS, Kan KJ, Willemsen G, Van Beijsterveldt CEM, … Boomsma DI (2015). Stability in symptoms of anxiety and depression as a function of genotype and environment: A longitudinal twin study from ages 3 to 63 years. Psychological Medicine, 45, 1039–1049. doi: 10.1017/S003329171400213X [PubMed: 25187475]

Östlund H, Keller E, & Hurd YL (2003). Estrogen receptor gene expression in relation to neuropsychiatric disorders. Annals of the New York Academy of Sciences, 1007, 54–63. [PubMed: 14993040]

Petrill SA, Lipton PA, Hewitt JK, Plomin R, Cherny SS, Corley R, & DeFries JC (2004). Genetic and environmental contributions to general cognitive ability through the first 16 years of life. Developmental Psychology, 40, 805–812. doi: 10.1037/0012-1649.40.5.805 [PubMed: 15355167]

Plomin R (1986). Multivariate analysis and developmental behavioral genetics: Developmental change as well as continuity. Behavior Genetics, 16, 25–43. [PubMed: 3707486]

Plomin R, DeFries JC, & Loehlin JC (1977). Genotype-environment interaction and correlation in the analysis of human behavior. Psychological Bulletin, 84, 309–322. doi: 10.1037/0033-2909.84.2.309 [PubMed: 557211]

Polderman TJ, Benyamin B, De Leeuw CA, Sullivan PF, Van Bochoven A, Visscher PM, & Posthuma D (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nature Genetics, 47(7), 702–709. [PubMed: 25985137]

Poromaa IS, & Gingnell M (2014). Menstrual cycle influence on cognitive function and emotion processing from a reproductive perspective. Frontiers in Neuroscience, 8, 1–16. doi: 10.3389/fnins.2014.00380 [PubMed: 24478622]

Purcell S (2002). Variance components models for gene–environment interaction in twin analysis. Twin Research, 5, 554–571. doi: 10.1375/twin.5.6.554 [PubMed: 12573187]

Roberts BW, & Mroczek DK (2008). Personality trait change in adulthood. Current Directions in Psychological Science, 17, 31–35. doi: 10.1016/j.pestbp.2011.02.012.Investigations [PubMed: 19756219]

Röcke C, Li SC, & Smith J (2009). Intraindividual variability in positive and negative affect over 45 days: Do older adults fluctuate less than young adults? Psychology and Aging, 24, 863–878. doi: 10.1037/a0016276 [PubMed: 20025402]

Romans S, Clarkson R, Einstein G, Petrovic M, & Stewart D (2012). Mood and the menstrual cycle: A review of prospective data studies. Gender Medicine, 9, 361–384. doi: 10.1016/j.genm.2012.07.003 [PubMed: 23036262]

Røysamb E, & Tambs K (2016). The beauty, logic and limitations of twin studies. Norsk Epidemiologi, 26, 35–46. doi: 10.5324/nje.v26i1-2.2014

Satorra A, & Bentler PM (2001). A scaled difference chi-square test statistic for moment structure analysis. Psychometrika, 66, 507–519.

Scarr S (1992). Developmental theories for the 1990s: Development and individual differences. Child Development, 63(1), 1–19. [PubMed: 1343618]

Scarr S, & McCartney K (1983). How people make their own environments: A theory of genotype-environment effects. Child Development, 54, 424–435. [PubMed: 6683622]

Sclove SL (1987). Application of model-selection criteria to some problems in multivariate analysis. Psychometrika, 52(3), 333–343.

Stone AA, Marco CA, Cruise CE, Cox DS, & Neale JM (1996). Are stress-induced immunological changes mediated by mood? A closer look at how both desirable and undesirable daily events influence sIgA antibody. International Journal of Behavioral Medicine, 3, 1–13. doi: 10.1207/s15327558ijbm0301_1 [PubMed: 16250763]

Tucker-Drob EM, & Briley DA (2014). Continuity of genetic and environmental influences on cognition across the life span: a meta-analysis of longitudinal twin and adoption studies. Psychological Bulletin, 140, 949–79. doi: 10.1037/a0035893 [PubMed: 24611582]

Turkheimer E, & Waldron M (2000). Nonshared environment: A theoretical, methodological, and quantitative review. Psychological Bulletin, 126, 78–108. doi: 10.1037/0033-2909.126.1.78 [PubMed: 10668351]

van den Berg SM, Beem L, & Boomsma DI (2006). Fitting genetic models using Markov Chain Monte Carlo algorithms with BUGS. Twin Research and Human Genetics, 9(3), 334–342. [PubMed: 16790144]

Wachs TD (1983). The use and abuse of environment in behavior-genetic research. Child Development, 54, 396–407. [PubMed: 6872631]

Wall EH, Hewitt SC, Case LK, Lin CY, Korach KS, & Teuscher C (2014). The role of genetics in estrogen responses: A critical piece of an intricate puzzle. FASEB Journal, 28, 5042–5054. doi: 10.1096/fj.14-260307 [PubMed: 25212221]

Watson D, Clark LA, & Tellegen A (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. Journal of Personality and Social Psychology, 54, 1063–1070. [PubMed: 3397865]

West MJ, & King AP (1987). Settling nature and nurture into an ontogenetic niche. Developmental Psychobiology, 20, 549–562. doi: 10.1002/dev.420200508 [PubMed: 3678619]

Winship C, & Korenman S (1999). Economic success and the evolution of schooling and mental ability. In Mayer S, & Peterson PE (Eds.), Earning and learning: How schools matter (pp. 49–78). Washington DC: Brookings.

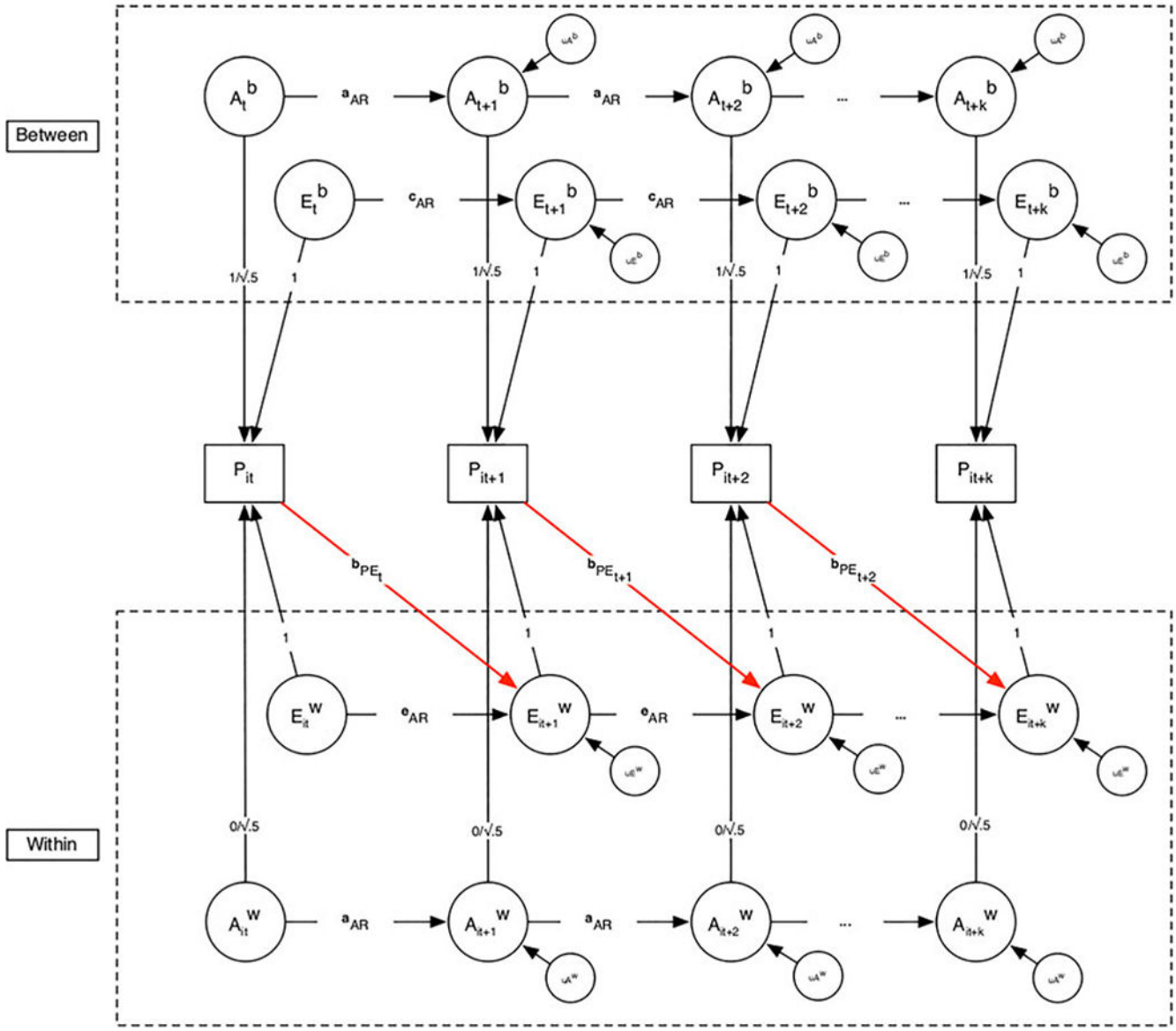

Figure 1.

Phenotype–environment (P→E) model. Biometric components of phenotypic scores for Twin *i* at Time *t*, P_{it}, are estimated between- and within-families; A_t^b = between-family genetic effect at time *t*; E_t^b = between-family (common) environmental effect at time *t*; A_t^w = within-family genetic effect at time *t*; E_t^w = within-family (nonshared) environmental effect at time *t*; uA_b^t = unique between-family genetic effect at time *t*; uE_b^t = unique between-family environmental effect at time *t*; uA_t^w = unique within-family genetic effect at time *t*; uE_t^w = unique within-family environmental effect at time *t*; a_{ar}, c_{ar}, and e_{ar} = autoregressive coefficient between adjacent components. The between-family and within-family genetic loadings for the monozygotic (MZ) twins are 1 and 0, respectively, to meet the assumption that MZ twins share 100% of their genes. The between-family and within-family genetic

loadings for the dizygotic (DZ) twins are both $\sqrt{.5}$ to meet the assumption that and DZ twins share 50%, on average, of their segregating genes. The red line represents the P→E parameter, b_{PE}, which was only estimated at the within-family level in the DZ group.

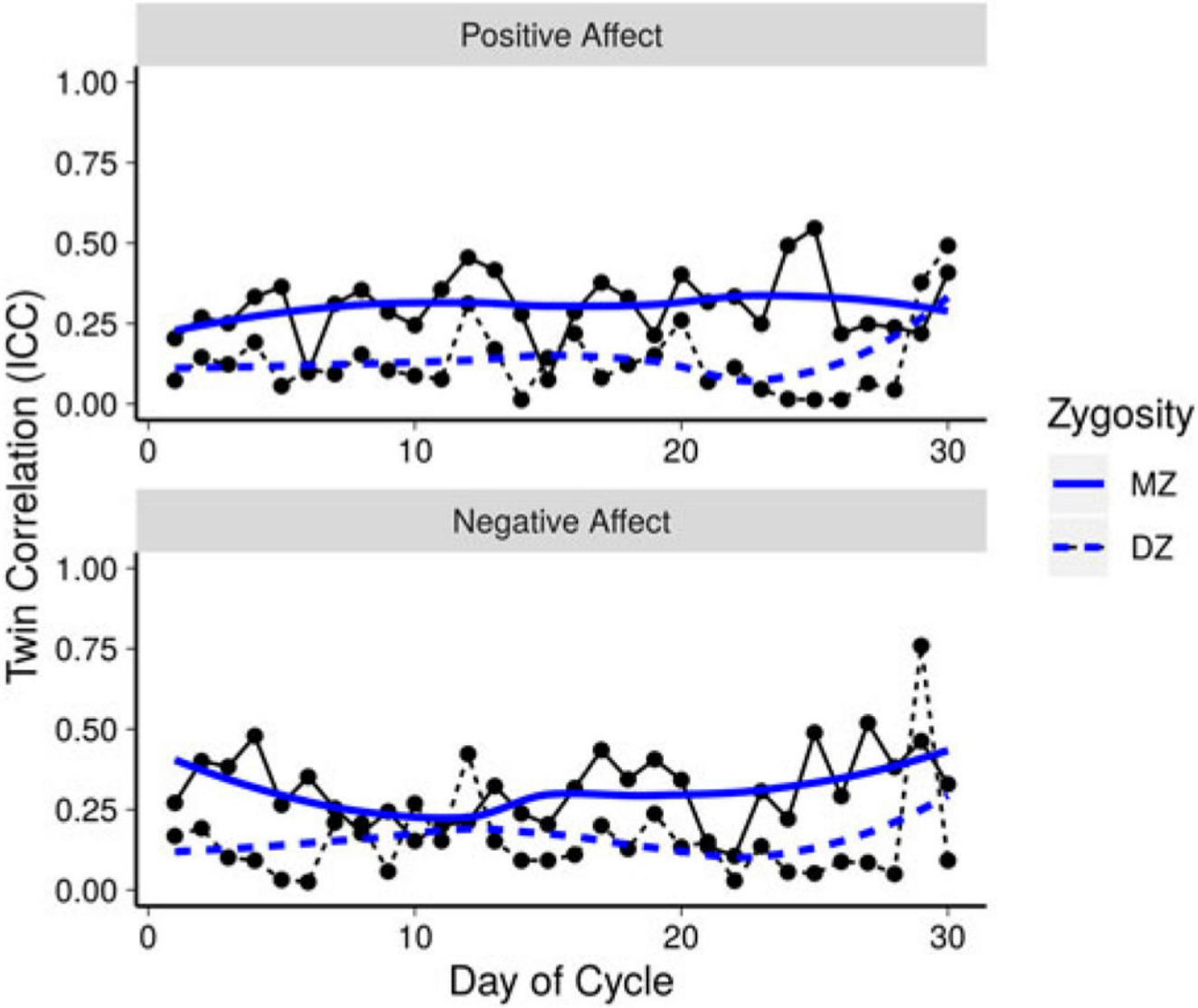

Figure 2.

Daily twin correlations of positive and negative affect scores. LOESS lines (in blue) are overlaid to illustrate general trends in twin similarity for each phenotype. On average, differences between monozygotic (MZ) and dizygotic (DZ) twin correlations are statistically significant across the 30 days (positive affect: $t = 6.10$, $df = 58$, $p < .001$; negative affect: $t = 4.04$, $df = 58$, $p < .001$).

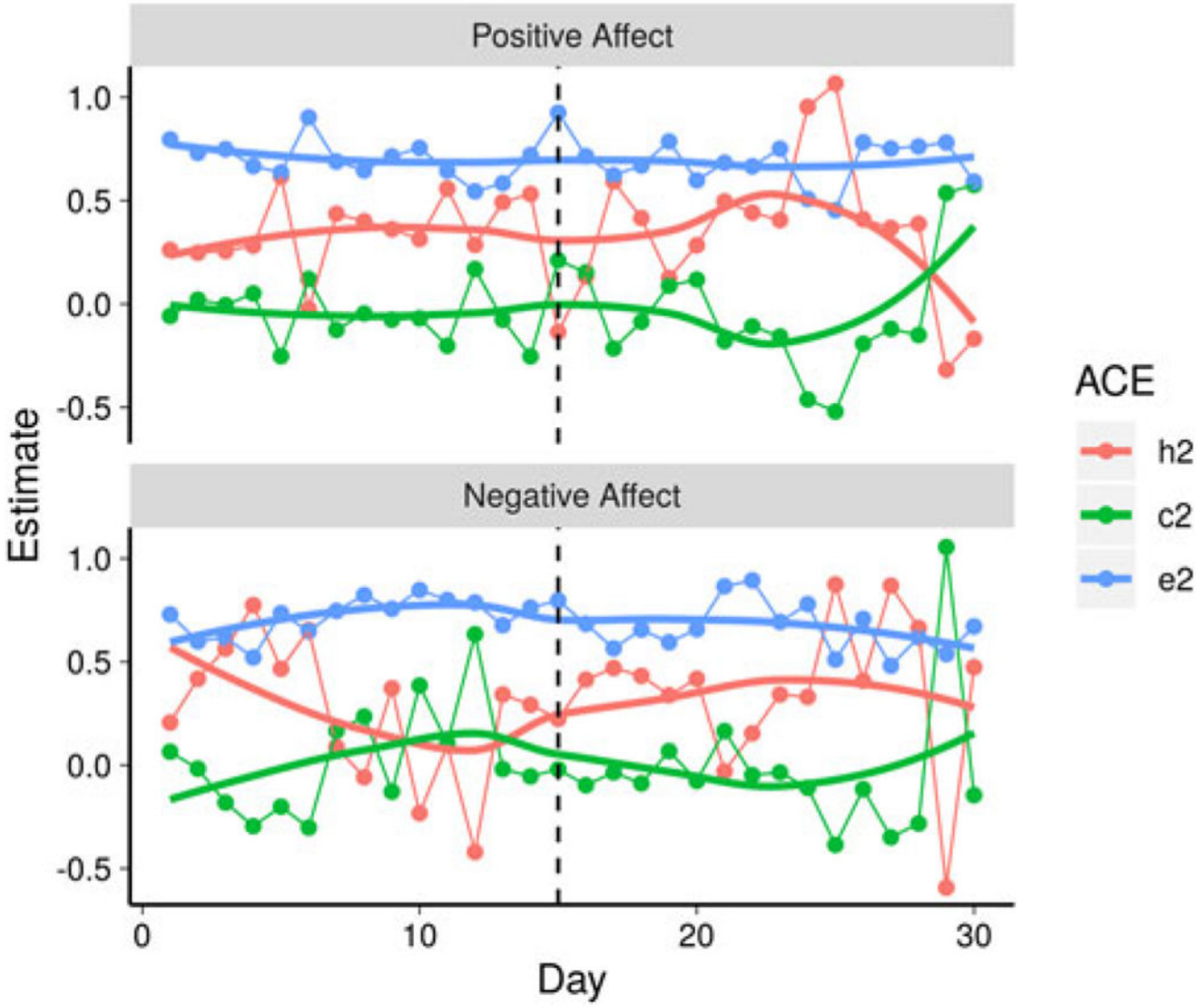

Figure 3.

Heritability and environment estimates of positive affect (top panel) and negative affect (bottom panel) by day. h^2 = heritability, which is the proportion of total variance in daily affect scores attributed to genetic variance; c^2 = shared environment, which is the proportion of total variance in daily affect scores attributed to shared environmental variance; e^2 = nonshared environment, which is the proportion of total variance in daily affect scores attributed to nonshared (unique) environmental variance. All estimates are based on classical univariate ACE models (genetic [A], shared environmental [C], and nonshared environmental [E]).

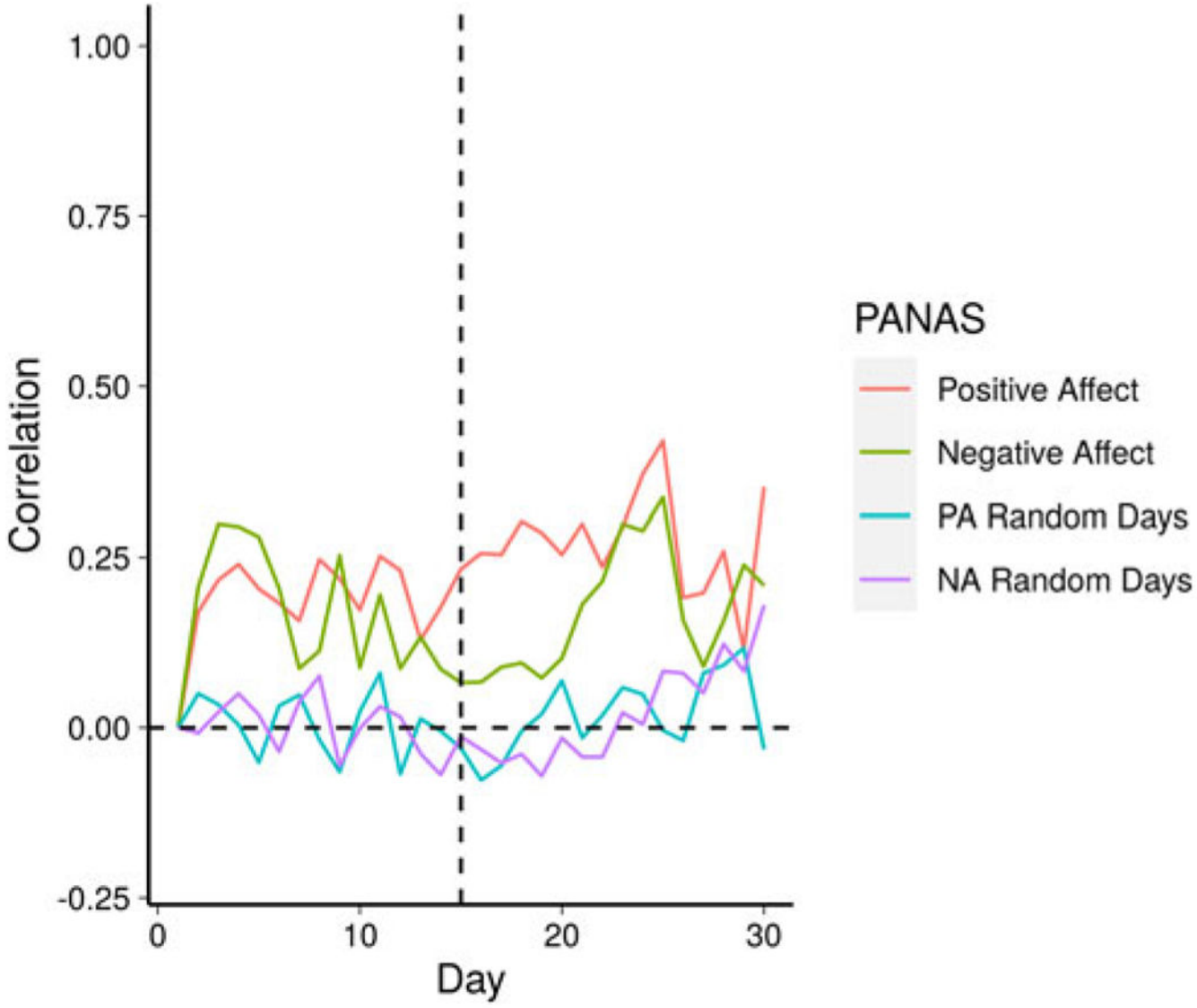

Figure 4.

Model estimated within-family *r*GE over 30 days for positive affect (PA) and negative affect (NA). *r*GE was re-estimated in P→E models where days were randomly ordered within twins to illustrate that *r*GE systematically changes across the 30 days.

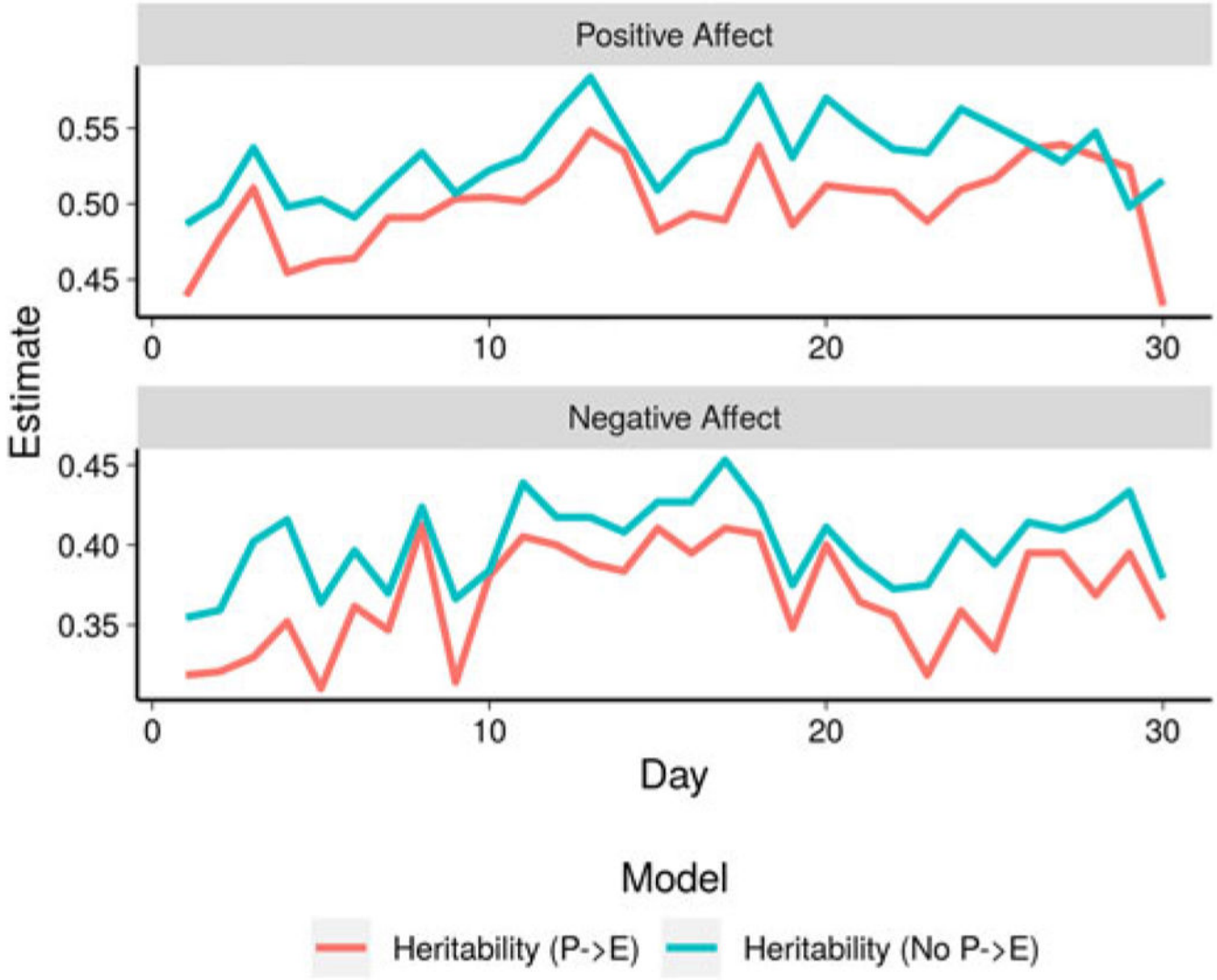

Figure 5.
Heritability estimates of positive affect (top) and negative affect (bottom) from phenotype–environment (P→E) model (red) and genetic simplex model (blue) across the 30 days.

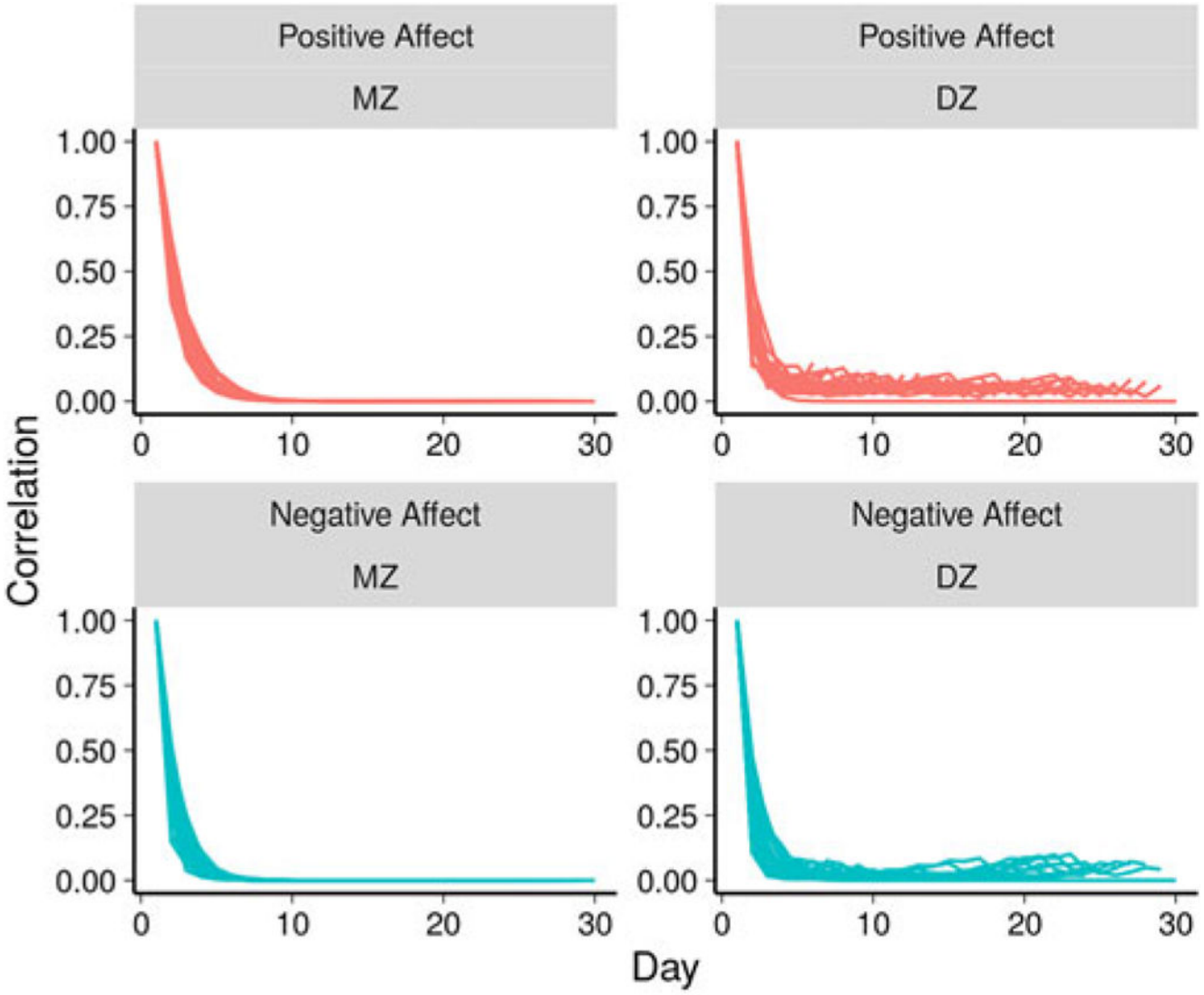

Figure 6.

Longitudinal correlations among nonshared environmental correlations across 30 days for positive affect (PA) and negative affect (NA). Model estimated MZ and DZ correlations are taken from phenotype–environment (P→E) models.

Table 1.

Model fitting results

Positive affect models	**LL**	***df***	**ΔLL**	**Δ*df***	***P***	**AIC**	**BIC**	**SSABIC**
1. P→E model	−8,861.32	1,811	-	-	-	17,960.63	18,447.23	18,069.58
2. P→E = E model	−8,947.93	1,840	86.61	29	0.000	18,075.85	18,443.87	18,158.25
Negative affect models	**LL**	***df***	**ΔLL**	**Δ*df***	***P***	**AIC**	**BIC**	**SSABIC**
1. P→E model	−5,540.21	1,811	-	-	-	11,318.41	11,805.01	11,427.36
2. P→E = E model	−5,593.70	1,840	53.49	29	0.004	11,367.40	11,735.41	11,449.79

Notes. Model 1 is the P→E model, which is the less restricted model that distinguishes between person-environment match via b_{PE} parameters and nonshared environmental effects over days. Model 2 restricts b_{PE} effects to be equal to nonshared environmental effects. LL = log likelihood value; ΔLL = likelihood ratio value; Δ*df* = differences in in degrees of freedom between models 1 and 2; AIC = Akaike information criterion; BIC = Bayesian information criterion; SSABIC = sample size adjusted Bayesian information criterion. P→E models fit better than ACE (genetic [A], shared environmental [C], and nonshared environmental [E]) Cholesky decomposition models. ACE Cholesky decomposition model fit for positive affect was: LL = −7,884.16 (*df* = 505), AIC = 18,618.33, BIC = 24,445.22, SSABIC = 19,922.94. For negative affect, model fit was: LL = −4,671.80 (*df* = 505), AIC = 12,193.61, BIC = 18,020.49, SSABIC = 13,498.22.