Archival Report

Biological Psychiatry: CNNI

Accurate Prediction of Momentary Cognition From Intensive Longitudinal Data

Zoë W. Hawks, Roger Strong, Laneé Jung, Emorie D. Beck, Eliza J. Passell, Elizabeth Grinspoon, Shifali Singh, Madelyn R. Frumkin, Martin Sliwinski, and Laura T. Germine

ABSTRACT

BACKGROUND: Deficits in cognitive performance are implicated in the development and maintenance of psychopathology. Emerging evidence further suggests that within-person fluctuations in cognitive performance may represent sensitive early markers of neuropsychiatric decline. Incorporating routine cognitive assessments into standard clinical care—to identify between-person differences and monitor within-person fluctuations—has the potential to improve diagnostic screening and treatment planning. In support of these goals, it is critical to understand to what extent cognitive performance varies under routine, remote assessment conditions (i.e., momentary cognition) in relation to a wide range of possible predictors.

METHODS: Using data-driven, high-dimensional methods, we ranked strong predictors of momentary cognition and evaluated out-of-sample predictive accuracy. Our approach leveraged innovations in digital technology, including ambulatory assessment of cognition and behavior 1) at scale (n = 122 participants, n = 94 females), 2) in naturalistic environments, and 3) within an intensive longitudinal study design (mean = 25.5 assessments/participant).

RESULTS: Reaction time ($R^2 > 0.70$) and accuracy ($0.56 > R^2 > 0.35$) were strongly predicted by age, between-person differences in mean performance, and time of day. Effects of self-reported, intraindividual fluctuations in environmental (e.g., noise) and internal (e.g., stress) states were also observed.

CONCLUSIONS: Our results provide robust estimates of effect size to characterize sources of cognitive variability, to support the identification of optimal windows for psychosocial interventions, and to possibly inform clinical evaluation under remote neuropsychological assessment conditions.

https://doi.org/10.1016/j.bpsc.2022.12.002

Cognitive performance refers to the speed and accuracy of responding on tasks designed to measure attention, memory, processing speed, perceptual reasoning, and executive functioning (1,2). Cognitive performance is associated with short- and long-term aspects of physical (e.g., cardiovascular disease), mental (e.g., suicidality), and public (e.g., accident proneness) health (3–5). Beyond simple associations, cognitive performance deficits are implicated in the development and maintenance of psychopathology (6–11), while improvements predict successful psychosocial and adaptive functioning (12). Emerging evidence further suggests that within-person fluctuations in cognitive performance represent sensitive early markers of neuropsychiatric decline (13). Against this backdrop, incorporating routine cognitive assessments into standard care—to identify between-person differences and monitor within-person fluctuations—has the potential to improve diagnostic screening and treatment planning. For example, such assessments may be used as population-level screeners to ascertain individuals at heightened risk for cognitive decline (14–16), or they may be used in treatment to support timely delivery of psychosocial interventions when patients are most likely to respond (17).

To be practicable for routine administration, cognitive assessments must be psychometrically sound under remote testing conditions, scalable, and ultrabrief (18). In recent years, research groups have validated digital cognitive assessments that meet these criteria (19,20), and best-practice guidelines for developing and evaluating digital cognitive assessments have been proposed (18,21). We refer to cognitive performance under routine, remote assessment conditions as momentary cognition. Hypothesis-driven research has rapidly advanced current understanding of whether, and to what extent, predictors of a priori interest explain variation in momentary cognition [for a review, see (22)]. However, caution is warranted when interpreting hypothesis-driven research in the absence of data-driven exploration. In genetics (pre-genomics), hypothesis-driven studies of candidate genes routinely identified significant yet inconsequential effects, while effects of greater magnitude were overlooked (23–26). Equally, variables with limited predictive utility in data-driven investigations may have reliable effects in hypothesis-driven investigations, which in turn may provide mechanistic insights that progress clinical science (27). Data- and hypothesis-driven approaches are mutually informative, and both are necessary to develop, refine, and test theories of cognitive variation.

To bridge the gap between current hypothesis-driven research and clinical translation, it is critical to evaluate

ISSN: 2451-9022

momentary cognition in relation to a wide range of predictors. The present work represents an initial step toward this end. Specifically, we used data-driven methods to identify meaningful predictors of momentary cognition. Our approach was informed by complementary research traditions (data-driven, high-dimensional; hypothesis-driven, longitudinal) that are commonly used to characterize cognitive performance. Below, we summarize strengths and limitations of these traditions to motivate current aims and methods.

Data-Driven, High-Dimensional Studies of Cognitive Performance

Data-driven, high-dimensional studies have principally examined cognitive performance in relation to biological (e.g., genetic, neural) predictors. In genetics, results suggest that polygenic predictors explain approximately 10% of the variation in cognition (28–30). Neuroimaging predictors may explain as much as 80% of the variation in aspects of cognition (31), although performance varies markedly across studies (10%–70%) (32–35), and it is unclear how neuroimaging modality, sample size, algorithm selection, and cognitive domain impact results. Collectively, this work demonstrates that data-driven methods have the potential to accurately predict cognitive performance, but major gaps remain. First, relative to biological predictors, self-report predictors are understudied and represent an important frontier for investigation (36,37). Self-report predictors are scalable, have applications for intervention, and can enrich understanding of genetic and neural biomarkers by suggesting behavioral mechanisms. Second, most data-driven work has been cross-sectional (38), producing group-level insights that may not apply to individuals (39,40) and precluding studies of within-person fluctuations (41,42). Finally, machine learning algorithms are frequently inscrutable (43), complicating efforts to understand how model inputs generate predictions.

Hypothesis-Driven, Longitudinal Studies of Momentary Cognition

Recent innovations in digital technology support cognitive and behavioral data collection remotely, reliably, and at scale (18,44), enabling intensive longitudinal investigation of momentary cognition. Ecological momentary assessment (EMA) is a popular intensive longitudinal design in which participants answer self-report questions about recent experiences multiple times each day (45). In cognitive variations of EMA, participants additionally complete brief cognitive tasks in naturalistic environments (18). Routinely sampling cognitive performance using EMA makes it possible to partition between-person (interindividual) and within-person (intraindividual) variation in momentary cognition (41). These distinctions are important; for example, sleep quality has been linked to working memory performance between, but not within, individuals (46). In the absence of data-driven cognitive EMA research, studies have frequently examined different predictors, making it difficult to compare results and increasing the likelihood that available results reflect topics of interest to researchers, rather than topics of central consequence to cognitive performance.

Study Aims and Approach

Combining the strengths of data-driven and intensive longitudinal studies, we aimed to 1) accurately predict momentary cognition in high-dimensional space and 2) identify meaningful predictors thereof (Figure 1). To this end, we evaluated more than 50 demographic (e.g., age), self-report (e.g., affect, stress, context, arousal), and mobile (e.g., time of day, date, screen size) features in relation to variation in momentary cognition. Baseline and EMA features are defined in S1 in the Supplement. Momentary cognition was parameterized with respect to domain (effortful attention, visuospatial capacity) and outcome metric (accuracy, reaction time [RT]) (Figure 2). Although there are multiple statistical techniques that partition interindividual and intraindividual variation in clustered (e.g., intensive longitudinal) data, these techniques are not commonly applied in high-dimensional space, where researchers often ignore clusters, represent clusters as categorical variables, or adopt a fully idiographic (one person, one model) approach. To partition interindividual and intraindividual variation in momentary cognition, we analyzed data using hierarchical Bayesian modeling with horseshoe priors (47). Our novel application of hierarchical Bayesian modeling flexibly accommodated high-dimensional, clustered data; accepted linear, nonlinear, and interaction terms; provided transparent feature rankings; and yielded accessible, interpretable, cross-validated estimates of performance (48).

METHODS AND MATERIALS

Participants

EMA data were collected from 122 adult participants recruited through links placed on the front page of our digital research platform, TestMyBrain.org (49). Participants older than 18 years were eligible to enroll if they did not report history of head trauma, current substance use disorder, neurological illness, major medical illness, or disability that would interfere with study protocol. Consent was obtained electronically before enrollment. Study procedures were approved by the Mass General Brigham Institutional Review Board. Participant characteristics are reported in Table 1.

Methods and Measures

Before EMA, participants provided information about demographic characteristics and sleep habits and completed virtual onboarding with task instructions, practice trials, and corrective feedback. During EMA, participants received 30 text messages (3 daily × 10 days) prompting them to complete questionnaires and cognitive assessments taking approximately 10 minutes. Each text corresponded to one EMA sitting, and texts arrived at random times within fixed time windows. More specifically, one text arrived within each of 3 daily windows: morning (9:00 AM–12:30 PM), afternoon (1:15–4:45 PM), and evening (5:30–9:00 PM). Time was measured in participants' local time zones. On receipt, participants had 30 minutes to respond to texts by clicking embedded links. After 30 minutes, links expired. Questionnaires were identical within and across sittings (same surveys presented in the same order), whereas cognitive assessments were identical within

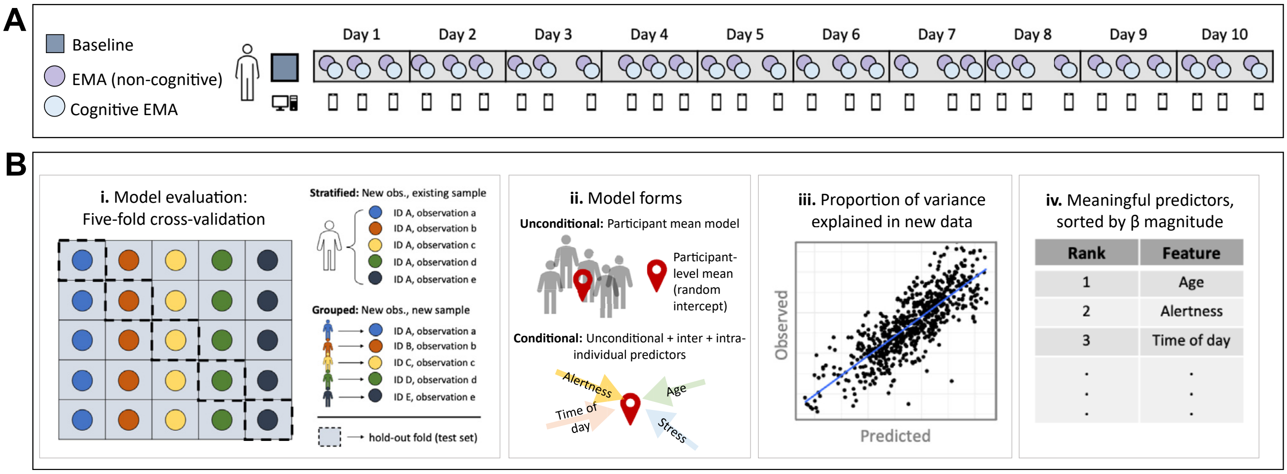

Figure 1. Schematic depicting data collection timeline and analytic plan. **(A)** Participants provided baseline data, including demographics and self-reported sleep habits, before ecological momentary assessment (EMA). During EMA, they received 30 texts (3 per day × 10 days) prompting them to complete brief self-report questionnaires and cognitive assessments on their personal smartphone devices. **(B)** i. Data were analyzed within 5-fold cross-validation. Models were evaluated based on their ability to predict new observations from individuals used in training (stratified cross-validation) as well as new observations from individuals who were not used in training (grouped cross-validation). ii. Models were built in stages (unconditional/participant mean model vs. conditional) to partition sources of interindividual and intraindividual variance. Unconditional models included participant-level (random) intercepts, which indexed interindividual differences in mean cognition. Conditional models additionally included baseline and noncognitive EMA variables. iii. Model performance was evaluated with respect to the proportion of variance explained in new data, reflecting the strength of the association between predicted and observed values. iv. Variables that contributed most strongly to prediction of momentary cognition were identified by ranking coefficients, with consideration for uncertainty around estimates.

sittings (across participants) and randomized across sittings (different test versions presented in different orders).

Ecological Momentary Assessment. Items were adapted from existing instruments to measure constructs of potential relevance to cognition, including positive and negative affect, stress, social functioning, context, and attention (50–55). Items, responses, variable abbreviations, and distributions are provided in S1 in the Supplement.

Cognitive EMA. EMA items were analyzed in relation to RT and accuracy on cognitive EMA tasks measuring effortful/sustained attention (Choice Reaction Time, Gradual Onset Continuous Performance Test) (1,56–58) and visuospatial memory capacity (Multiple Object Tracking, Digit-Symbol Matching) (59–63). These distinctions follow established taxonomies (60). Cognitive descriptive statistics are provided in Table 2, task descriptions are provided in S2.1 in the Supplement, and multilevel reliability is reported in S2.2 in the Supplement (19,64). Between-person reliability, reflecting the proportion of variance in scores attributable to differences between individuals, equaled or exceeded 0.94. Within-person reliability, reflecting the proportion of systematic variance across (relative to within) measurement occasions, ranged from 0.27 to 0.75 (19). Within-person reliability estimates are consistent with prior research (19,65,66) and isolate variance that may be explained by time-varying predictors.

Statistical Analysis

Data cleaning maximized the signal-to-noise ratio by excluding observations that were strongly suggestive of careless responding. Details are provided in S2.3 in the Supplement. To be included in the analysis sample, participants were required to provide clean data in ≥20 of 30 EMA sittings. Participants included in analyses (n = 122) did not differ from the full sample (N = 202 participants who provided demographic information) with respect to age, gender, race, ethnicity, or educational attainment ($ps > .05$).

Analyses were performed in R using tidyverse and brms (48,67,68). Hierarchical modeling accounts for shared variance in nested data (69), and Bayesian modeling incorporates background information in the form of prior distributions to inform estimation (70). These approaches were well suited to the present study because they enabled feature evaluation and ranking within a multilevel framework.

Preprocessing included grand mean centering, scaling, and correlation-based feature reduction. During feature reduction, 3 highly correlated social functioning variables ($rs > 0.9$) were averaged to create a composite. Models were implemented separately for cognitive outcomes and built in stages (unconditional, conditional) to partition interindividual and intraindividual variance in momentary cognition. Unconditional models included participant-level random intercepts indexing interindividual differences in momentary cognition. We refer to unconditional models as participant mean models. Conditional models additionally included linear (see S1 in the Supplement), quadratic (age), cyclic (sine, cosine of time) (71), and 2-way interaction (linear age × cyclic time) fixed effects. Change in variance explained (ΔR^2) between participant mean and conditional models reflected the extent to which fixed effects improved estimates of momentary cognition, principally by explaining variation in intraindividual cognitive fluctuations. Model formulas and syntax are provided in S2.4 and S2.5 in the Supplement.

Performance was evaluated using 5-fold cross-validation (CV) with stratified and grouped partitions. In stratified CV, participants contribute data to all folds, and R^2 reflects generalizability to new observations from existing individuals included in training.

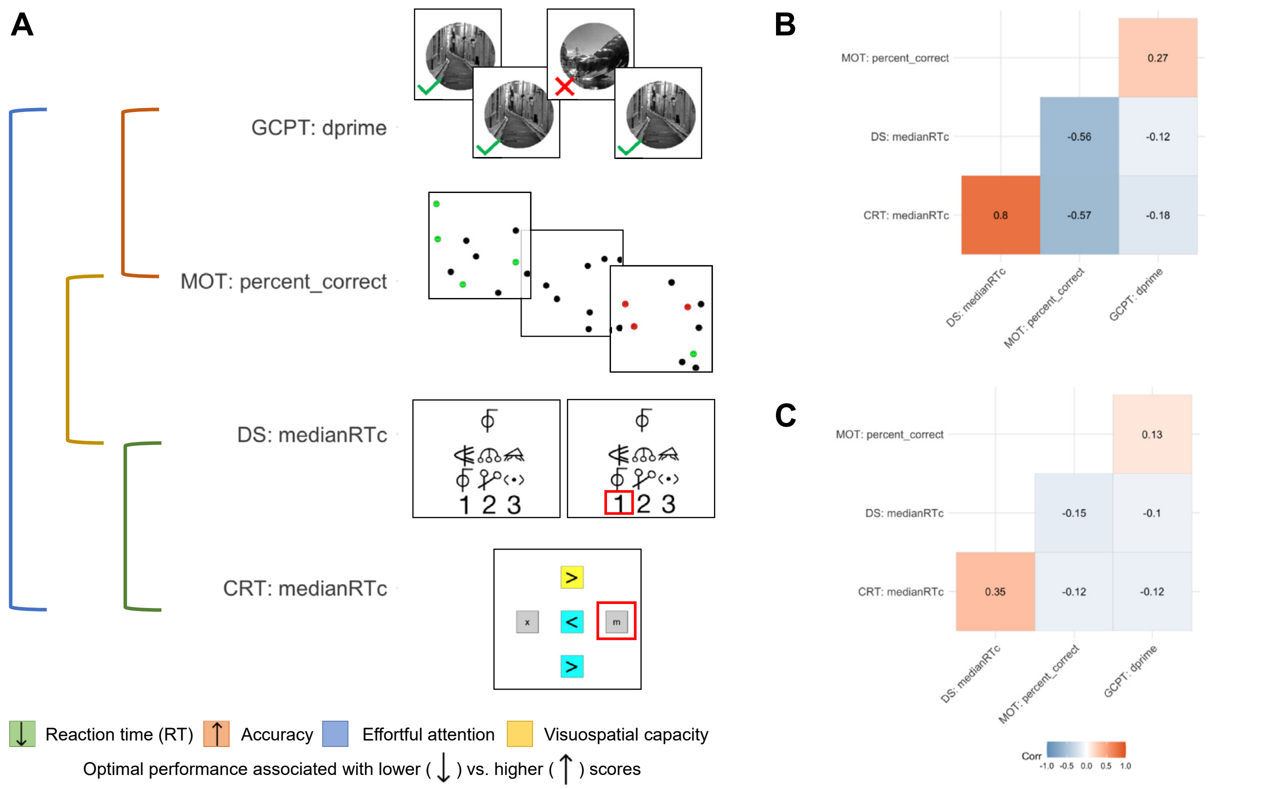

Figure 2. **(A)** Cognition was evaluated in relation to domain (effortful attention, visuospatial capacity) and metric (accuracy, reaction time [RT]) using 4 tasks and associated performance indices. Task [performance index] = gradual onset continuous performance test [dprime], multiple object tracking [percent correct], digit-symbol matching [median RT], and choice RT [median RT]. Tasks are depicted in panel **(A)**, and task instructions are elaborated in Methods and Materials. **(B)** Between-person correlations of average momentary cognition and **(C)** average within-person correlations of momentary cognition. Between-person correlations were stronger than within-person correlations. Negative correlations between RT and accuracy reflect the fact that faster responding was associated with greater accuracy. corr, correlation; CRT, choice reaction time; DS, digit-symbol matching; GCPT, gradual onset continuous performance test; MOT, multiple object tracking; RTc, median reaction time on correct responses.

In grouped CV, approximately 20% of participants contribute data to a given fold, and R^2 reflects generalizability to new observations from new individuals excluded from training. Grouped prediction was implemented using brms defaults.

Conditional models partitioned interindividual and intraindividual variation in momentary cognition. Post hoc analyses further disaggregated interindividual and intraindividual influences on momentary cognition. First, we separated time-varying, self-report variables[1] into interindividual and intraindividual predictors. Interindividual predictors were computed by centering and scaling person-level mean scores. Intraindividual predictors were computed by centering and scaling deviations from participant means. Next, we used hierarchical linear modeling to evaluate associations between disaggregated (interindividual and intraindividual) predictors and momentary cognition. Regression coefficients for interindividual predictors reflect between-person associations, and regression coefficients for intraindividual predictors reflect within-person associations (72). A random intercept was estimated for each participant, and feature space was restricted to variables that were significant in data-driven analyses. For DS_medianRTc, alertness and context were rescaled between 0 and 1 (“rs”) and combined to reduce multicollinearity:

$$alert\_composite = (PA\_alert\_rs - alert\_sleepiness\_rs)/2$$

$$context\_composite = (context\_diff\_concentrating\_rs + context\_going\_on\_rs)/2$$

RESULTS

Proportion of Variance Explained

Model performance was evaluated within stratified and grouped CV (Figure 1B). In stratified CV, performance was evaluated on new observations, allowing us to generalize results to new data from an existing sample. In grouped CV, performance was evaluated in new individuals, allowing us to generalize results to new data from a new sample. Accurate performance ($R^2 > 70\%$) was obtained in stratified, conditional models predicting RT from baseline and EMA variables. Stratified CV consistently explained more variance than

[1]We specify time-varying because baseline variables (e.g., age) do not index intraindividual variation, and we specify self-report because time-varying, experimentally determined predictors (e.g., test position within battery) do not index meaningful interindividual variation.

Table 1. Sample Characteristics

Characteristic	*n* (%) or Mean (SD) [Range]
Education	
College	36 (29.5%)
Graduate school	19 (15.6%)
High school	13 (10.7%)
Masters	21 (17.2%)
Some college	25 (20.5%)
Technical school	8 (6.6%)
Primary Language	
English	114 (93.4%)
Not English	8 (6.6%)
Race	
Asian	8 (6.6%)
Black or African American	5 (4.1%)
Multiracial	4 (3.3%)
White	94 (77%)
Not sure	7 (5.7%)
Rather not say	4 (3.3%)
Gender	
Female	94 (77%)
Male	22 (18%)
Nonbinary	6 (4.9%)
Ethnicity	
Hispanic	12 (9.8%)
Not Hispanic	107 (87.7%)
Not sure	2 (1.6%)
Rather not say	1 (0.8%)
Age, Years	36.62 (14.29) [18–70]
EMAs Completed	25.48 (2.18) [20–30]

EMA, ecological momentary assessment.

grouped CV, as expected given that grouped models could not use participant-specific estimates of average cognition (i.e., random intercepts) during prediction. Conditional models consistently explained more variance than participant mean models (see S2.6 in the Supplement), suggesting that intraindividual cognitive fluctuations were predicted, in part, by model fixed effects.

In stratified CV, conditional models predicting RT from baseline and EMA variables explained 73% of the variance in visuospatial capacity and 71% of the variance in effortful attention, while conditional models predicting accuracy from baseline and EMA variables explained 55% of the variance in visuospatial capacity and 35% of the variance in effortful attention (Table 3 and Figure 3). Stratified participant mean models explained slightly less variation in RT and accuracy (ΔR^2 = 2.5%–5.2%). In grouped CV, conditional models predicting RT from baseline and EMA data explained 42.3% of the variance in visuospatial capacity and 26.8% of the variance in effortful attention, while conditional models predicting accuracy from baseline and EMA data explained 14.2% of the variance in visuospatial capacity and 3.3% of the variance in effortful attention. Grouped participant mean models explained negligible variance in RT and accuracy (R^2 = 0.7%–2.3%).

Table 2. Cognitive Descriptive Statistics and ICCs

Variable	Mean (SD)	Range	ICC	No. of Participants (T)
CRT RT	709.36 (120.56)	491–1312.05	0.68	122 (3059)
DS RT	810.56 (148.87)	545.15–1621.5	0.71	122 (3062)
GCPT dprime	2.85 (0.74)	0.04–4.23	0.35	122 (3050)
MOT Percent Correct	0.74 (0.11)	0.37–1	0.55	121 (2955)

ICCs were estimated in unconditional models using all available data and reflect the proportion of variance attributable to interindividual differences.

CRT, choice reaction time; DS, digit symbol; GCPT, gradual onset continuous performance test; ICC, intraclass correlation coefficient; MOT, multiple object tracking; RT, reaction time; T, number of ecological momentary assessments completed.

Meaningful Predictors of Variance

Having established the predictive utility conditional model predictors, we next examined the relative magnitude of their effects. Effects are interpretable at mean levels of covariates, and horseshoe priors shrunk coefficients toward zero, reducing the impact of multicollinearity, guarding against overfitting (73), and supporting feature selection. Figure 4 visualizes significant and marginal effects (for which 95% and 80%–90% credible intervals did not include zero, respectively), and S2.7 in the Supplement provides model statistics. Statistics were estimated using all available data, although they were highly similar in each fold of stratified (S2.8 in the Supplement) and grouped (S2.9 in the Supplement) prediction. We focus our discussion on significant effects, but illustrate marginal effects to support future hypothesis generation.

Reaction Time. Across cognitive domains, RT was positively related to age and self-reported sleepiness, with slower RTs among older individuals and among individuals reporting greater sleepiness. In contrast, RT was negatively related to test position and study days, with faster RTs in tests that occurred later in the assessment battery and later in the study. Time of day was consistently related to momentary cognition, but the nature of circadian effects varied. For example, on effortful attention measures, participants had greater accuracy in the morning relative to the evening, yet were faster in the evening relative to the morning (Figure 5). Domain-specific effects for RT were also observed. Quadratic effects of age were observed for effortful attention, indicating that associations between age and RT strengthened as age increased. Alertness was negatively associated with visuospatial capacity, indicating that RT slowed as alertness decreased. Self-reported contextual distractions significantly predicted slower RTs in visuospatial capacity tasks, while self-reported attention significantly predicted slower RTs in effortful attention tasks. Higher ratings of effort were associated with faster RTs for visuospatial capacity but not effortful attention.

Accuracy. Across cognitive domains, less accurate performance was observed when tests were interrupted and/or completed in busy, noisy environments. Self-reported affect and stress were associated with visuospatial capacity and effortful attention, but effects were in opposite directions: Higher positive

Table 3. Model Performance (RMSE and R^2) Predicting RT and Accuracy in Tasks Assessing Visuospatial Capacity and Effortful Attention

CV Split	Outcome Metric	Cognitive Domain	Model Stage	Cognitive Variable	RMSE (L, U)	R^2 (L, U)	ΔR^2
Stratified	RT	Visuospatial capacity	Conditional	DS median RTc	0.53 (0.46, 0.56)	72.51 (68.69, 77.53)	3.71
			Participant mean	DS median RTc	0.56 (0.50, 0.58)	68.80 (65.56, 72.58)	
		Effortful attention	Conditional	CRT median RTc	0.54 (0.52, 0.57)	70.51 (66.31, 73.67)	5.20
			Participant mean	CRT median RTc	0.59 (0.57, 0.60)	65.31 (61.89, 67.74)	
	Accuracy	Visuospatial capacity	Conditional	MOT percent correct	0.67 (0.65, 0.69)	55.24 (50.76, 58.57)	2.53
			Participant mean	MOT percent correct	0.69 (0.68, 0.71)	52.71 (46.83, 56.41)	
		Effortful attention	Conditional	GCPT dprime	0.81 (0.77, 0.84)	35.00 (30.95, 38.2)	3.72
			Participant mean	GCPT dprime	0.83 (0.80, 0.87)	31.28 (28.86, 33.67)	
Grouped	RT	Visuospatial capacity	Conditional	DS median RTc	0.76 (0.61, 0.84)	42.25 (33.54, 49.59)	39.94
			Participant mean	DS median RTc	1.00 (0.82, 1.13)	2.31 (0.44, 9.56)	
		Effortful attention	Conditional	CRT medianRTc	0.86 (0.76, 0.97)	26.84 (16.92, 40.05)	25.71
			Participant mean	CRT median RTc	1.00 (0.89, 1.14)	1.13 (0.09, 4.96)	
	Accuracy	Visuospatial capacity	Conditional	MOT percent correct	0.95 (0.84, 1.01)	14.18 (1.45, 20.04)	12.60
			Participant mean	MOT percent correct	1.00 (0.92, 1.10)	1.58 (0.08, 6.84)	
		Effortful attention	Conditional	GCPT dprime	0.99 (0.93, 1.05)	3.32 (0.88, 6.01)	2.65
			Participant mean	GCPT dprime	1.01 (0.94, 1.08)	0.67 (0.01, 1.99)	

Stratified splits estimated model performance on new observations, and grouped splits estimated model performance on new observations in new individuals. Unconditional models included participant-level (random) intercepts, and conditional models additionally included baseline and ecological momentary assessment variables. R^2 is the proportion of variance explained and is expressed as a percentage, and ΔR^2 is the change in R^2 between conditional and unconditional models, also expressed as a percentage.

CRT, choice reaction time; CV, cross validation; DS, digit symbol; GCPT, gradual onset continuous performance test; L, lower bound of CV estimates; MOT, multiple object tracking; RMSE, root mean square error; RT, reaction time; RTc, median reaction time on correct responses; U, upper bound of CV estimates.

affect (determination) was associated with better effortful attention, whereas higher stress (following an argument) was associated with improved visuospatial capacity. Domain-specific effects for accuracy were also observed. Study days were positively associated with visuospatial capacity but not effortful attention, indicating improved visuospatial capacity later in the study. Age was negatively associated with visuospatial capacity but not effortful attention, indicating poorer visuospatial capacity among older adults. Finally, time of day was associated with effortful attention but not visuospatial capacity, indicating improved effortful attention in the morning compared with afternoon (Figure 5).

Post Hoc Analyses. In data-driven analyses, predictors were grand mean centered to balance research goals (accurate prediction, feature selection) and computational complexity. Under such conditions, regression coefficients conflate interindividual and intraindividual effects (72). Post hoc analysis demonstrated that 1) intraindividual fluctuations in alertness, attention, context, affect, and stress predicted RT and accuracy and 2) interindividual differences in alertness predicted visuospatial RT (see S2.10 in the Supplement). Whereas chronically elevated alertness predicted slower visuospatial RT (interindividual effect), within-person elevations in alertness predicted faster visuospatial RT (intraindividual effect).

DISCUSSION

Data-Driven, High-Dimensional Methods Supported Accurate Prediction of Momentary Cognition

The present study used hierarchical Bayesian modeling with horseshoe priors to predict momentary cognition across domains (effortful attention, visuospatial capacity) and outcomes (accuracy, RT). Conditional models predicting new observations from existing individuals accounted for more than 70% of the variance in RT and 35% to 55% of the variance in accuracy. ΔR^2 between conditional and participant mean models was small but nonetheless significant, suggesting that participant mean performance is a strong predictor of momentary cognition, and maximizing prediction of momentary cognition requires modeling between- and within-person predictors. In conditional RT models, momentary cognition was predicted by age, time of day, sleepiness/arousal, practice effects, and context. In conditional accuracy models, momentary cognition was predicted by interruptions, context, and affect. When predicting new observations from new individuals, models

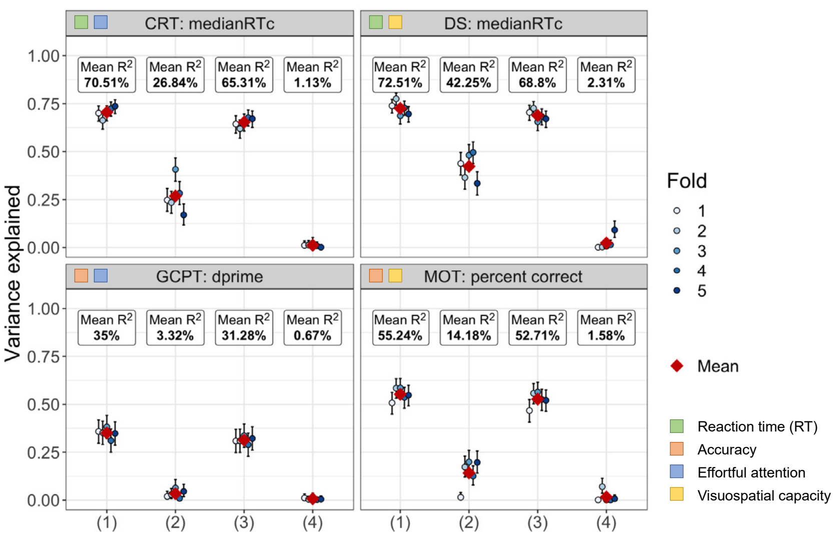

Figure 3. Variance explained in momentary cognition across models (conditional vs. participant mean) and cross-validation splits (stratified vs. grouped). Stratified splits were used to estimate model performance on new observations, while grouped splits were used to estimate model performance on new observations in new individuals. Participant mean models included participant-level (random) intercepts, while conditional models additionally included baseline and noncognitive ecological momentary assessment variables. x-axis: (1) conditional model, new observations; (2) conditional model, new individuals; (3) participant mean model, new observations; (4) participant mean model, new individuals. When applied to new observations, conditional models accounted for more than 70% of the variation in reaction time (RT) and 35%–55% of the variation in accuracy. ΔR^2 between conditional and participant mean models was comparatively small (~5%) but nonetheless significant, suggesting that participant mean performance is a strong predictor of momentary cognition and maximizing prediction of momentary cognition requires modeling between-person and within-person predictors. Shades of blue illustrate variability among folds. Error bars reflect 95% confidence intervals around R^2 within folds. CRT, choice reaction time; DS, digit-symbol matching; GCPT, gradual onset continuous performance test; MOT, multiple object tracking; RTc, median reaction time on correct responses.

could not use information about between-person differences in average cognition and performed considerably worse.

Environmental, Circadian, and Internal State Variables Constituted Shared Sources of Cognitive Variation

Momentary cognition was slower and less accurate when assessments were completed in environments perceived to be noisy or busy, when participants reported difficulty concentrating before assessment, and when participants reported that one or more interruptions occurred during assessment. RT was less impacted by interruptions than accuracy, plausibly because RT (operationalized within sessions with respect to median) was robust to trial-level outliers. Results suggest that failure to measure and appropriately control for contextual and environmental variables may confound clinical studies of momentary cognition, particularly insofar as environmental variables (such as noise) and sociodemographic characteristics are themselves correlated (74). An alternative possibility is that difficulty concentrating disposes participants to notice environmental noises and distractions. Collecting subjective (self-report) and objective (decibel level) indices of noise (75) will support future research aimed at clarifying how

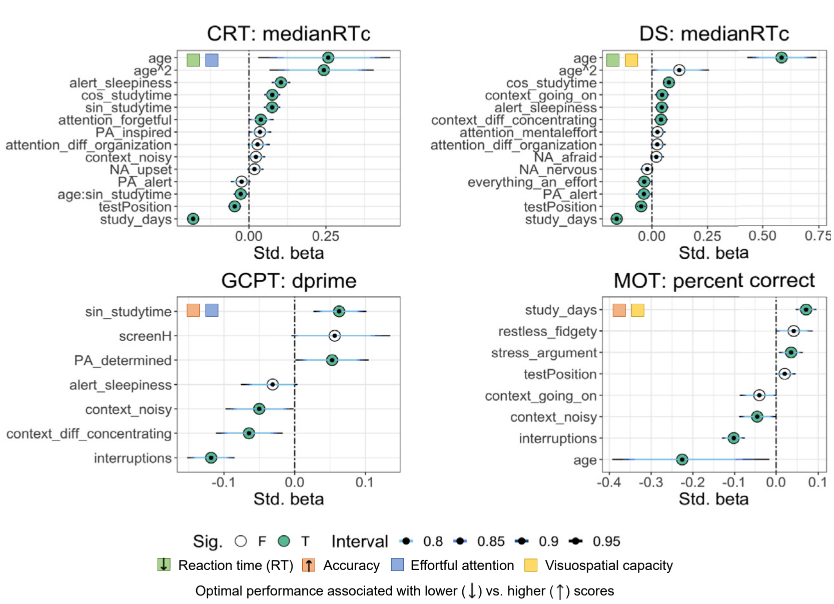

Figure 4. Strong predictors, for which 80% credible intervals did not include zero, sorted by cognitive outcome and coefficient magnitude. Significant predictors, for which 95% credible intervals did not include zero, are identified by a sea-green dot. In addition to linear and quadratic effects of age, time-varying environmental (e.g., noise, interruptions), circadian (e.g., time of day, arousal), and affective (e.g., determination) variables were associated with momentary cognition across domains (effortful attention, visuospatial capacity) and outcome metrics (reaction time, accuracy). CRT, choice reaction time; DS, digit-symbol matching; GCPT, gradual onset continuous performance test; MOT, multiple object tracking; RTc, median reaction time on correct responses; Sig., significant; Std. beta, standardized β.

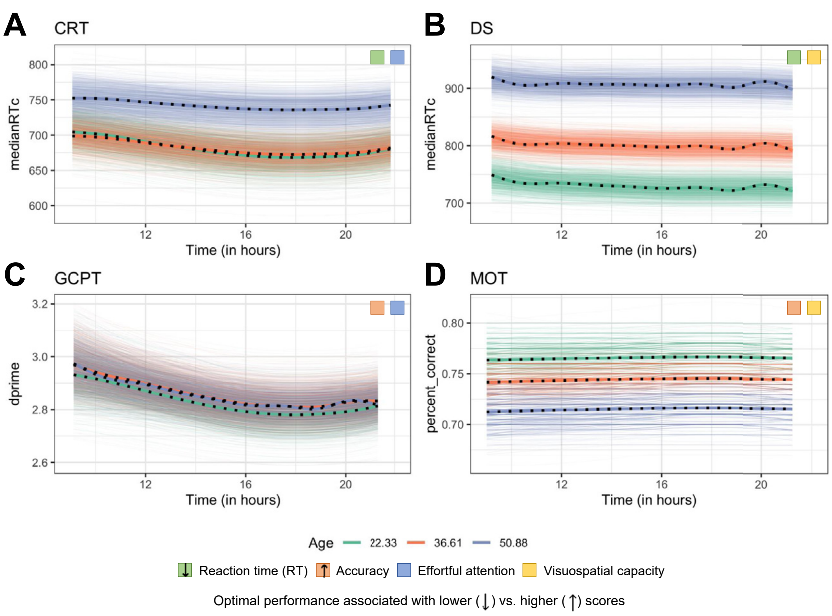

Figure 5. (A–D) Age (in years), time of day (in hours), and age × time of day interaction effects. Black dots represent fit curves. Thin lines represent draws from the expectation of the posterior predictive distribution, illustrating uncertainty around fit curves. Age is plotted at the sample mean ± 1 SD. Time of day effects suggest a possible speed-accuracy tradeoff, with accurate but slow effortful attention in the morning and error-prone but fast effortful attention in the evening **(A, C)**. Circadian slowing of effortful attention was less pronounced among older compared with younger adults **(A)**, and time of day effects were less pronounced for visuospatial capacity compared with effortful attention. CRT, choice reaction time; DS, digit-symbol matching; GCPT, gradual onset continuous performance test; MOT, multiple object tracking; RTc, median reaction time on correct responses.

distractions impact cognition in daily life. Additionally, indices of noise may be used during remote neuropsychological assessment to statistically control for distractions.

Circadian factors, including time of day and self-reported arousal (alertness, sleepiness), also emerged as strong predictors of momentary cognition. Time of day effects suggest a possible speed-accuracy tradeoff, with accurate but slow effortful attention in the morning and error-prone but fast effortful attention in the evening (Figure 5A, C). Circadian slowing of effortful attention was less pronounced among older compared with younger adults, and time of day effects were less pronounced for visuospatial capacity compared with effortful attention. The latter is consistent with evidence that sleep impacts effortful attention interindividually and intraindividually (76,77), but has mixed results in other cognitive domains (78–80). Effortful attention and visuospatial capacity were faster and more accurate when participants reported high alertness and low sleepiness. Effects were driven by intraindividual fluctuations rather than interindividual differences, although both intraindividual and interindividual effects were observed for visuospatial capacity. Whereas elevations in alertness relative to one's own mean predicted faster responses on tests of visuospatial capacity, elevations in alertness relative to means of others predicted slower responses (see S2.10 in the Supplement). Processing speed impairments are well documented in clinical disorders characterized by chronic alertness (e.g., posttraumatic stress disorder) (81). The present study extends these findings to a nonclinical sample and further suggests that intraindividual associations are reversed.

Intraindividual fluctuations in affect and stress predicted momentary cognition, although these effects were weaker than environmental and circadian effects. Aspects of positive affect (determination) predicted improvements in effortful attention, and aspects of negative affect (feeling like everything is an effort) predicted faster responses on tests of visuospatial capacity. The latter may reflect associations between perceived and actual effort in nonclinical populations. In addition to affective predictors, interpersonal stress following an argument was associated with improved visuospatial capacity. The available literature on intraindividual associations between affect, stress, and cognition is small and mixed. Neubauer *et al.* (82) observed that negative affect predicted momentary working memory impairment in children, whereas von Stumm (83) failed to observe associations between affect and momentary cognition. Hyun *et al.* (84) observed associations between stress and cognition, similar to the present study, but they used different constructs (anticipatory stress, working memory) and their effects were in the opposite direction (stress predicted cognitive errors). Weak, inconsistent, and variable results may be expected if intraindividual associations are true only of some people, moderated by interindividual variables such as stress sensitivity (82), or susceptible to methodological differences. The impact of methodological differences is appreciated interindividually; for example, the direction and magnitude of associations between stress and cognition commonly vary as a function of stress intensity, origin, task, and processing demands (85). Measuring and elucidating the impact of such variables on intraindividual associations represents an important direction for research.

Methodological and Clinical Implications

Research utilizing ambulatory assessment to measure momentary cognition—at scale, in naturalistic environments, and within intensive longitudinal designs—is in its infancy. Our approach combined the strengths of data-driven and intensive longitudinal studies to identify meaningful predictors of momentary cognition, providing foundational insights in an emerging research space.

With respect to research planning and design, results provide robust estimates of effect size to support power computation across data- and hypothesis-driven studies. Whereas future data-driven studies may leverage information about the relative magnitude of within- versus between-person fluctuations in momentary cognition, future hypothesis-driven studies may use information about the relative strength of individual within- and between-person predictors. Questions remain regarding the extent to which passively acquired, quantitatively assessed variables (ambient noise, temperature, location, physical activity) further improve prediction.

With respect to intervention, the data-driven, multilevel approach that we describe readily extends to new populations. For example, the same model could be applied to diverse, clinical populations to test the generalizability of results. Future research establishing generalizability is necessary before clinical translation. Within the present sample, results suggest 4 promising applications for tailoring interventions. First, results may guide neuropsychologists by identifying factors that impact cognitive performance under remote assessment conditions. Second, they may aid in the development of cognitive interventions by identifying modifiable treatment targets—for example, determination, which predicted momentary cognition in the present study and (bearing replication in clinical samples) may be targeted in therapy using empirically supported protocols (86). Third, they may support delivery of psychosocial interventions at moments when patients (e.g., with cognitive impairment) are most likely receptive. Finally, to the extent that RT remains predictable over long periods of time (e.g., years), results suggest avenues for ascertaining individuals who deviate from expected trajectories, improving risk monitoring in aging (14–16). Research examining intraindividual cognitive variability in relation to clinical outcomes has commonly operationalized variability across cognitive domains within a single time point (87,88), whereas we operationalized variability within a single cognitive domain across time points. These operational definitions measure change on different time scales, presuppose different study designs (single session vs. intensive longitudinal), and are suited to different applications (prediction vs. monitoring). It remains unclear which cognitive outcomes will be sensitive to clinical decline in the context of intensive longitudinal monitoring.

Limitations

To our knowledge, this work represents the first data-driven, high-dimensional investigation of momentary cognition. There are, however, several limitations. Participants were predominantly female and White/non-Hispanic, limiting generalizability. The sample was not clinically ascertained, and it is unclear how results may change in the context of psychopathology or cognitive decline. Our study protocol lasted 10 days, limiting insights into intraindividual stability. Longitudinal burst designs (16,89) are necessary to demonstrate that RT is stable within individuals over extended periods, with potential implications for risk monitoring in aging. Finally, we lacked enough time points within participants to build individualized models of momentary cognition. We observed that contextual, circadian, and internal state variables constituted shared sources of cognitive variation. However, there may be interindividual differences in magnitude, strength, shape, and moderators of intraindividual associations. To clarify complex sources of cognitive variation, future studies may collect longer time series for idiographic modeling and/or leverage more flexible forms of predictive modeling with explainable artificial intelligence. Notably, these approaches compromise simultaneous estimation of interindividual and intraindividual effects.

Conclusions

Leveraging innovations in digital technology and statistical computing, we performed a data-driven, high-dimensional analysis of interindividual and intraindividual predictors of momentary cognition. RT was highly predictable from intraindividual, age, and circadian variables. Accuracy was predicted by similar variables, but substantial variation remains to be explained. To this end, passive sensing and idiographic modeling represent promising future directions. More immediately, results from the present study may support the identification of optimal windows for psychosocial interventions, enhance clinical understanding under remote neuropsychological assessment conditions, and inform the development of cognitive interventions by identifying highly impactful, potentially modifiable treatment targets.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the National Institutes of Health (Grant Nos. R01 DK121240 and R01 MH121617 [to LTG], Grant No. F32 MH127812 [to ZWH], and Grant No. F31 MH124291 [to MRF]) and Brain and Behavior Research Foundation (NARSAD Young Investigator Award [to LTG]).

We thank the research participants for their contributions to this study.

Links to cognitive ecological momentary assessment tasks are available on GitHub (https://github.com/zwihawks/PredictingMomentaryCog; refer to CogEMA_tasks.txt). Data analyzed for the current study are available on GitHub (https://github.com/zwihawks/PredictingMomentaryCog; refer to cleanData.rds). Code to analyze and visualize data is available on GitHub (https://github.com/zwihawks/PredictingMomentaryCog; refer to BayesianPrediction.Rmd and VisualizeResults.R).

LTG has received compensation as a member of the Scientific Advisory Board of Sage Bionetworks, a 501c3 nonprofit organization. LTG is on the Board of Directors of the Many Brains Project, a 501c3 nonprofit organization. ZWH has received consulting fees from Blueprint Health. All other authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Institute for Technology in Psychiatry, McLean Hospital, Belmont, Massachusetts (ZWH, RS, LJ, EJP, EG, SS, LTG); Department of Psychiatry, Harvard Medical School, Cambridge, Massachusetts (ZWH, RS, SS, LTG); Department of Psychology, University of California, Davis, Davis, California (EDB); Department of Psychological & Brain Sciences, Washington University in St. Louis, St. Louis, Missouri (MRF); and Department of Human Development and Family Studies, Pennsylvania State University, University Park, Pennsylvania (MS).

Address correspondence to Zoë W. Hawks, Ph.D., at zhawks@mclean.harvard.edu.

Received Oct 14, 2022; revised Nov 8, 2022; accepted Dec 6, 2022.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsc.2022.12.002.

REFERENCES

1. Rutter LA, Vahia IV, Forester BP, Ressler KJ, Germine L (2020): Heterogeneous indicators of cognitive performance and performance variability across the lifespan. Front Aging Neurosci 12:62.
2. Singh S, Strong RW, Jung L, Li FH, Grinspoon L, Scheuer LS, *et al.* (2021): The TestMyBrain Digital Neuropsychology Toolkit: Development and psychometric characteristics. J Clin Exp Neuropsychol 43:786–795.

3. Kreitler S, Weissler K, Barak F (2013): Physical health and cognition. In: Kreitler S, editor. Cognition and Motivation: Forging an Interdisciplinary Perspective. New York: Cambridge University Press, 238–269.
4. Lawton R, Parker D (1998): Individual differences in accident liability: A review and integrative approach. Hum Factors 40:655–671.
5. Ayalon L, Mackin S, Arean PA, Chen H, McDonel Herr EC (2007): The role of cognitive functioning and distress in suicidal ideation in older adults. J Am Geriatr Soc 55:1090–1094.
6. Millan MJ, Agid Y, Brüne M, Bullmore ET, Carter CS, Clayton NS, *et al.* (2012): Cognitive dysfunction in psychiatric disorders: Characteristics, causes and the quest for improved therapy. Nat Rev Drug Discov 11:141–168.
7. Pessoa L (2008): On the relationship between emotion and cognition. Nat Rev Neurosci 9:148–158.
8. Chakrabarty T, Hadjipavlou G, Lam RW (2016): Cognitive Dysfunction in Major Depressive Disorder: Assessment, Impact, and Management. Focus (Am Psychiatr Publ) 14:194–206.
9. Bora E, Harrison BJ, Yücel M, Pantelis C (2013): Cognitive impairment in euthymic major depressive disorder: a meta-analysis. Psychol Med 43:2017–2026.
10. Mohlman J, Gorman JM (2005): The role of executive functioning in CBT: A pilot study with anxious older adults. Behav Res Ther 43:447–465.
11. Lam RW, Kennedy SH, McIntyre RS, Khullar A (2014): Cognitive dysfunction in major depressive disorder: Effects on psychosocial functioning and implications for treatment. Can J Psychiatry 59:649–654.
12. Sahdra BK, MacLean KA, Ferrer E, Shaver PR, Rosenberg EL, Jacobs TL, *et al.* (2011): Enhanced response inhibition during intensive meditation training predicts improvements in self-reported adaptive socioemotional functioning. Emotion 11:299–312.
13. Rutter LA, Vahia IV, Passell E, Forester BP, Germine L (2021): The role of intraindividual cognitive variability in posttraumatic stress syndromes and cognitive aging: A literature search and proposed research agenda. Int Psychogeriatr 33:677–687.
14. Salthouse TA, Nesselroade JR, Berish DE (2006): Short-term variability in cognitive performance and the calibration of longitudinal change. J Gerontol B Psychol Sci Soc Sci 61:144–151.
15. Dixon RA, Garrett DD, Lentz TL, MacDonald SWS, Strauss E, Hultsch DF (2007): Neurocognitive markers of cognitive impairment: Exploring the roles of speed and inconsistency. Neuropsychology 21:381–399.
16. Sliwinski M, Hoffman L, Hofer S (2010): Modeling retest and aging effects in a measurement burst design. In: Molenaar PCM, Newell KM, editors. Individual Pathways of Change: Statistical Models for Analyzing Learning and Development. Washington, DC: American Psychological Association, 37–50.
17. Smyth JM, Heron KE (2016): Is providing mobile interventions "just-in-time" helpful? An experimental proof of concept study of just-in-time intervention for stress management. 2016 IEEE Wireless Health (WH), 1–7.
18. Germine L, Strong R, Singh S, Sliwinski MJ (2021): Toward dynamic phenotypes and the scalable measurement of human behavior. Neuropsychopharmacology 46:209–216.
19. Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, Lipton RB (2018): Reliability and validity of ambulatory cognitive assessments. Assessment 25:14–30.
20. Nicosia J, Aschenbrenner AJ, Balota DA, Sliwinski MJ, Tahan M, Adams S, *et al.* (2023): Unsupervised high-frequency smartphone-based cognitive assessments are reliable, valid, and feasible in older adults at risk for Alzheimer disease. J Int Neuropsychol Soc 29:459–471.
21. Germine L, Reinecke K, Chaytor NS (2019): Digital neuropsychology: Challenges and opportunities at the intersection of science and software. Clin Neuropsychol 33:271–286.
22. Weizenbaum E, Torous J, Fulford D (2020): Cognition in context: Understanding the everyday predictors of cognitive performance in a New Era of measurement. JMIR mHealth uHealth 8:e14328.
23. Collins AL, Kim Y, Sklar P, O'Donovan MC, Sullivan PF (2012): Hypothesis-driven candidate genes for schizophrenia compared to genome-wide association results. Psychol Med 42:607–616.
24. Bosker FJ, Hartman CA, Nolte IM, Prins BP, Terpstra P, Posthuma D, *et al.* (2011): Poor replication of candidate genes for major depressive disorder using genome-wide association data. Mol Psychiatry 16:516–532.
25. De Vries B, Anttila V, Freilinger T, Wessman M, Kaunisto MA, Kallela M, *et al.* (2016): Systematic re-evaluation of genes from candidate gene association studies in migraine using a large genome-wide association data set. Cephalalgia 36:604–614.
26. Kell DB, Oliver SG (2004): Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. Bioessays 26:99–105.
27. Prentice DA, Miller DT (2016): When small effects are impressive. In: Kazdin AE, editor. Methodological Issues and Strategies in Clinical Research. Washington, DC: American Psychological Association, 99–105.
28. Carey C, Strong R, Huang Y, Aslibekyan S, Gentleman R, Smoller J, *et al.* (2021): Shared and distinct genetic influences between cognitive domains and psychiatric disorder risk based on genome-wide data. Biol Psychiatry 89(suppl):S45–S46.
29. Allegrini AG, Selzam S, Rimfeld K, von Stumm S, Pingault JB, Plomin R (2019): Genomic prediction of cognitive traits in childhood and adolescence. Mol Psychiatry 24:819–827.
30. Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, *et al.* (2018): Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 50:1112–1121.
31. Schütze M, de Souza Costa D, de Paula JJ, Malloy-Diniz LF, Malamut C, Mamede M, *et al.* (2018): Use of machine learning to predict cognitive performance based on brain metabolism in neurofibromatosis type 1. PLoS One 13:e0203520.
32. Cui Z, Su M, Li L, Shu H, Gong G (2018): Individualized prediction of reading comprehension ability using gray matter volume. Cereb Cortex 28:1656–1672.
33. Rosenberg MD, Finn ES, Scheinost D, Papademetris X, Shen X, Constable RT, *et al.* (2015): A neuromarker of sustained attention from whole-brain functional connectivity. Nat Neurosci 19:165–171.
34. Greene AS, Gao S, Scheinost D, Constable RT (2018): Task-induced brain state manipulation improves prediction of individual traits. Nat Commun 9:2807.
35. Sui J, Jiang R, Bustillo J, Calhoun V (2020): Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: Methods and promises. Biol Psychiatry 88:818–828.
36. Dagum P (2018): Digital biomarkers of cognitive function. NPJ Digit Med 1:10.
37. Cavedoni S, Chirico A, Pedroli E, Cipresso P, Riva G (2020): Digital biomarkers for the early detection of mild cognitive impairment: Artificial intelligence meets virtual reality. Front Hum Neurosci 14:245.
38. Bzdok D, Meyer-Lindenberg A (2018): Machine learning for precision psychiatry: Opportunities and challenges. Biol Psychiatry Cogn Neurosci Neuroimaging 3:223–230.
39. Mõttus R (2021): How correlations can (not) be applied to individual people: A tutorial for researchers, students and the public. PsyArXiv. https://doi.org/10.31234/osf.io/bpm9y.
40. Molenaar PCM (2004): A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. Measurement: Interdisciplinary Research and Perspectives 2:201–218.
41. Hamaker EL (2012): Why researchers should think "within-person": A paradigmatic rationale. In: Mehl MR, Conner TS, editors. Handbook of Research Methods for Studying Daily Life. New York: Guilford Press, 43–61.
42. Kanning MK, Ebner-Priemer UW, Schlicht WM (2013): How to investigate within-subject associations between physical activity and momentary affective states in everyday life: A position statement based on a literature overview. Front Psychol 4:1–16.
43. de Laat PB (2018): Algorithmic decision-making based on machine learning from big data: Can transparency restore accountability? Philos Technol 31:525–541.
44. Arnold KD, Rohn KC (2020): Internet-enabled experience sampling: Innovations in signaling and data sources. In: Kennedy E, Qian Y, editors. Advancing Educational Research With Emerging Technology. Hershey, PA: IGI Global, 124–151.
45. Junghaenel DU, Stone AA (2020): Ecological momentary assessment for the psychosocial study of health. In: Sweeny K, Robbins ML,

Cohen LM, editors. The Wiley Encyclopedia of Health Psychology. Hoboken, NJ: John Wiley & Sons, 105–112.
46. Derby CA, Zhaoyang R, Jiao JL, Sliwinski MJ, Buxton OM (2020): Within and between person effects of sleep quality on daily cognitive performance: Data from the Einstein Aging Study (EAS). Alzheimers Dement 16:e044565.
47. van Erp S, Oberski DL, Mulder J (2019): Shrinkage priors for Bayesian penalized regression. J Math Psychol 89:31–50.
48. Bürkner PC (2017): brms: An R package for Bayesian multilevel models using Stan. J Stat Softw 80:1–28.
49. Germine L, Nakayama K, Duchaine BC, Chabris CF, Chatterjee G, Wilmer JB (2012): Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. Psychon Bull Rev 19:847–857.
50. Mackinnon A, Jorm AF, Christensen H, Korten AE, Jacomb PA, Rodgers B (1999): A short form of the Positive and Negative Affect Schedule: Evaluation of factorial validity and invariance across demographic variables in a community sample. Pers Individ Dif 27:405–416.
51. Almeida DM, Wethington E, Kessler RC (2002): The daily inventory of stressful events: An interview-based approach for measuring daily stressors. Assessment 9:41–55.
52. Stawski RS, Sliwinski MJ, Almeida DM, Smyth JM (2008): Reported exposure and emotional reactivity to daily stressors: The roles of adult age and global perceived stress. Psychol Aging 23:52–61.
53. Granholm E, Ben-Zeev D, Fulford D, Swendsen J (2013): Ecological momentary assessment of social functioning in schizophrenia: Impact of performance appraisals and affect on social interactions. Schizophr Res 145:120–124.
54. Sliwinski MJ, Almeida DM, Smyth J, Stawski RS (2009): Intraindividual change and variability in daily stress processes: Findings from two measurement-burst diary studies. Psychol Aging 24:828–840.
55. Åkerstedt T, Anund A, Axelsson J, Kecklund G (2014): Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. J Sleep Res 23:242–254.
56. Germine LT, Joormann J, Passell E, Rutter LA, Scheuer L, Martini P, *et al.* (2022): Neurocognition after motor vehicle collision and adverse post-traumatic neuropsychiatric sequelae within 8 weeks: Initial findings from the AURORA study. J Affect Disord 298:57–67.
57. Esterman M, Noonan SK, Rosenberg M, Degutis J (2013): In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. Cereb Cortex 23:2712–2723.
58. Fortenbaugh FC, Degutis J, Germine L, Wilmer JB, Grosso M, Russo K, *et al.* (2015): Sustained attention across the life span in a sample of 10, 000: Dissociating ability and strategy. Psychol Sci 26:1497–1510.
59. Storm RW, Pylyshyn ZW (1988): Tracking multiple independent targets: Evidence for a parallel tracking mechanism. Spat Vis 3:179–197.
60. Treviño M, Zhu X, Lu YY, Scheuer LS, Passell E, Huang GC, *et al.* (2021): How do we measure attention? Using factor analysis to establish construct validity of neuropsychological tests. Cogn Res Princ Implic 6:1–26.
61. Wechsler D (1997): WAIS-III Administration and Scoring Manual. San Antonio, TX: The Psychological Corporation.
62. Chaytor NS, Barbosa-Leiker C, Germine LT, Fonseca LM, McPherson SM, Tuttle KR (2021): Construct validity, ecological validity and acceptance of self-administered online neuropsychological assessment in adults. Clin Neuropsychol 35:148–164.
63. Hartshorne JK, Germine LT (2015): When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. Psychol Sci 26:433–443.
64. Raykov T, Marcoulides GA (2006): On multilevel model reliability estimation from the perspective of structural equation modeling. Struct Equ Modeling 13:130–141.
65. Brose A, Lövdén M, Schmiedek F (2014): Daily fluctuations in positive affect positively co-vary with working memory performance. Emotion 14:1–6.
66. Sliwinski MJ, Smyth JM, Hofer SM, Stawski RS (2006): Intraindividual coupling of daily stress and cognition. Psychol Aging 21:545–557.
67. R Core Team (2021): R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
68. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, *et al.* (2019): Welcome to the Tidyverse. Journal of Open Source Software 4:1686.
69. Woltman H, Feldstain A, MacKay C, Rocchi M (2012): An introduction to hierarchical linear modeling. Tutorials in Quantitative Methods for Psychology 8:52–69.
70. Tso IF, Taylor SF, Johnson TD (2021): Applying hierarchical bayesian modeling to experimental psychopathology data: An introduction and tutorial. J Abnorm Psychol 130:923–936.
71. Van De Maat R, Lataster J, Verboon P (2020): Why and how to deal with diurnal cyclic patterns in ambulatory assessment of emotions: A practical guide and discussion. Eur J Psychol Assess 36:471–481.
72. Hamaker EL, Muthén B (2020): The fixed versus random effects debate and how it relates to centering in multilevel modeling. Psychol Methods 25:365–379.
73. Altelbany S (2021): Evaluation of ridge, elastic net and lasso regression methods in precedence of multicollinearity problem: A simulation study. J Appl Econ Bus Stud 5:131–142.
74. Huang YK, Mitchell UA, Conroy LM, Jones RM (2021): Community daytime noise pollution and socioeconomic differences in Chicago, IL. PLoS One 16:e0254762.
75. Zamora W, Calafate CT, Cano JC, Manzoni P (2017): Accurate ambient noise assessment using smartphones. Sensors (Switzerland) 17:917.
76. Liu J, Ramakrishnan S, Laxminarayan S, Balkin TJ, Reifman J (2017): Real-time individualization of the unified model of performance. J Sleep Res 26:820–831.
77. Bei B, Wiley JF, Allen NB, Trinder J (2015): A cognitive vulnerability model of sleep and mood in adolescents under naturalistically restricted and extended sleep opportunities. Sleep 38:453–461.
78. Winslow BD, Nguyen N, Venta KE (2017): Improved mental acuity forecasting with an individualized quantitative sleep model. Front Neurol 8:160.
79. Lo JC, Groeger JA, Cheng GH, Dijk DJ, Chee MWL (2016): Self-reported sleep duration and cognitive performance in older adults: A systematic review and meta-analysis. Sleep Med 17:87–98.
80. Short MA, Blunden S, Rigney G, Matricciani L, Coussens S, Reynolds CM, *et al.* (2018): Cognition and objectively measured sleep duration in children: A systematic review and meta-analysis. Sleep Health 4:292–300.
81. O'Neil ME, Laman-Maharg B, Schnurr PP, Carlson KF, Twamley EW, Peterson C, *et al.* (2019): Objective cognitive impairment and subjective cognitive problems in veterans initiating psychotherapy for posttraumatic stress disorder: An exploratory study. Appl Neuropsychol Adult 26:247–254.
82. Neubauer AB, Dirk J, Schmiedek F (2019): Momentary working memory performance is coupled with different dimensions of affect for different children: A mixture model analysis of ambulatory assessment data. Dev Psychol 55:754–766.
83. von Stumm S (2016): Is day-to-day variability in cognitive function coupled with day-to-day variability in affect? Intelligence 55:1–6.
84. Hyun J, Sliwinski MJ, Smyth JM (2019): Waking up on the wrong side of the bed: The effects of stress anticipation on working memory in daily life. J Gerontol B Psychol Sci Soc Sci 74:38–46.
85. Sandi C (2013). Stress and cognition, 4 245–261). Wiley Interdiscip Rev Cogn Sci, 245–261.
86. Kyrios M (2008): Cognitive-behaviour therapy in medical illness. In: Simos G, editor. Cognitive Behaviour Therapy: A Guide for the Practising Clinician, vol. 2. London: Routledge, 139–155.
87. Reichenberg A, Weiser M, Rapp MA, Rabinowitz J, Caspi A, Schmeidler J, *et al.* (2006): Premorbid intra-individual variability in intellectual performance and risk for schizophrenia: A population-based study. Schizophr Res 85:49–57.
88. Bangen KJ, Weigand AJ, Thomas KR, Delano-Wood L, Clark LR, Eppig J, *et al.* (2019): Cognitive dispersion is a sensitive marker for early neurodegenerative changes and functional decline in nondemented older adults. Neuropsychology 33:599–608.
89. Nesselroade JR (1991): The warp and the woof of the developmental fabric. In: Downs RM, Liben LS, Palermo DS, editors. Visions of Aesthetics, the Environment & Development: The Legacy of Joachim F. Wohlwill. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., 213–240