PNAS RESEARCH ARTICLE | PSYCHOLOGICAL AND COGNITIVE SCIENCES

Automating the analysis of facial emotion expression dynamics: A computational framework and application in psychotic disorders

Nathan T. Hall[a,1], Michael N. Hallquist[a], Elizabeth A. Martin[b], Wenxuan Lian[c], Katherine G. Jonas[c], and Roman Kotov[c]

Edited by Timothy Wilson, University of Virginia, Charlottesville, VA; received August 9, 2023; accepted January 18, 2024

Facial emotion expressions play a central role in interpersonal interactions; these displays are used to predict and influence the behavior of others. Despite their importance, quantifying and analyzing the dynamics of brief facial emotion expressions remains an understudied methodological challenge. Here, we present a method that leverages machine learning and network modeling to assess the dynamics of facial expressions. Using video recordings of clinical interviews, we demonstrate the utility of this approach in a sample of 96 people diagnosed with psychotic disorders and 116 never-psychotic adults. Participants diagnosed with schizophrenia tended to move from neutral expressions to uncommon expressions (e.g., fear, surprise), whereas participants diagnosed with other psychoses (e.g., mood disorders with psychosis) moved toward expressions of sadness. This method has broad applications to the study of normal and altered expressions of emotion and can be integrated with telemedicine to improve psychiatric assessment and treatment.

facial emotion | facial expression analysis | emotion dynamics | psychosis | network model

Significance

Social communication in humans depends on the movement of facial muscles, which may be the most potent form of nonverbal communication regarding emotion. Quantification of how facial expressions change over time has many potential applications in psychiatry and medicine. Our study introduces a computational framework designed to measure the dynamics of facial emotion expressions, using a cohort of individuals diagnosed with psychotic disorders and never-psychotic adults as an empirical demonstration. We found patterns of emotion expression dynamics that were associated with different dimensions of psychosis, demonstrating the utility of the framework.

Altered experience and expression of emotion are involved in many psychiatric, neurologic, and medical conditions (1–9). These emotional abnormalities are associated with poor outcomes, but assessments are often costly and time-intensive, and available treatments are limited in their effectiveness (10–12). Moreover, our understanding of human emotions is based primarily on individuals' self-report at a single time point. Although valuable, subjective reports are prone to retrospective biases (13, 14) and cross-sectional studies miss the dynamic interplay among emotions essential to understanding how emotions and their expression in social encounters unfold in real time (15). While there is no agreed-upon method for measuring emotional experience on fast timescales (second-to-second), facial emotional expressions (EEs) are directly observable. Recent advances in computer vision can readily quantify EEs (16), opening the door for the development and validation of computational methods that provide insight into the dynamic expression of emotions.

EEs serve as adaptive tools in social communication to influence the behaviors and perceptions of others (17); monitoring this continuously changing stream of expressions allows for interactants to form (sometimes accurate sometimes not) representations of their partner's internal state. However, no studies have attempted to capture and interpret these rapid changes in EE. The current study presents an analytic method for addressing this gap, with potential applications to a range of medical and psychiatric pathologies. We combine facial emotion recognition algorithms (FERAs) trained on video data with computational network models to uncover EE dynamics on fast time scales. We validate this approach in a sample of people diagnosed with psychotic disorders, a population with pronounced EE abnormalities (10, 18, 19). Our approach provides clinically interpretable indices of EE dynamics that can be readily obtained with widely available video data in medical settings such as telehealth visits.

Our proposed framework is motivated by the fact that facial emotion expressions are essential in interpersonal communication (20). Although there is some cultural variation in the expression of social emotions (e.g., guilt), basic emotions (e.g., happiness, sadness) are universally displayed and recognized (21, 22). The current standard for objective measurement of facial expressions is the Facial Action Coding System (FACS) (23) and its companion system, Emotion FACS. These approaches revolutionized the study of facial expressions by standardizing measurement, but they rely on time-consuming ratings made by extensively trained coders, making FACS difficult to implement broadly. Electromyography (EMG) is a complementary tool for measuring facial movements. Although sensitive to subtle facial movements, EMG is obtrusive (24) and may draw

Author affiliations: [a]Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599; [b]Department of Psychological Science, University of California, Irvine, CA 92697; and [c]Department of Psychiatry, Stony Brook University, Stoney Brook, NY 11794

Author contributions: K.G.J. and R.K. designed research; N.T.H., M.N.H., W.L., K.G.J., and R.K. performed research; N.T.H. contributed new reagents/analytic tools; N.T.H. and M.N.H. analyzed data; and N.T.H., M.N.H., E.A.M., K.G.J., and R.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: natehall@unc.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2313665121/-/DCSupplemental.

Published March 26, 2024.

participants' attention toward their face, making them aware of changes in their expressions and limiting ecological validity.

These limitations can be addressed by advances in machine learning/AI (25), namely FERAs; such as FaceReader (26) that decode EE state from video recordings (16, 27, 28). FERAs are remarkably accurate and far more scalable than human ratings or EMG. Scientists are starting to deploy these algorithms to study emotional expression in medical conditions (9, 25, 29–33). However, much of EE research has relied on static indices, such as the average level of an expression over time, even though these abnormalities often manifest in temporal fluctuations within and between EEs. Emotion dynamics rely on the interaction of two fundamental forces: 1) inertia—the tendency to remain in the current emotion and resist change (34) and 2) transition/change—the inter-emotion relationships that modulate responses to key external and internal stimuli (15, 35, 36). We conceptualized the components of our EE dynamics approach similarly, treating parameters in our model as reflecting inertia (autoregressive EE effects) and dynamics (cross-lagged EE effects) in EE patterns.

Uncovering these dynamics requires sophisticated statistical models such as those developed in computational neuroscience for understanding the dynamics of communication among brain regions (37). One influential algorithm, Group Iterative Multiple Model Estimation (GIMME), uses latent variable modeling combined with vector autoregressive (VAR) techniques to estimate directed relationships among time-varying signals (37). GIMME performs particularly well in heterogenous data, where group-level effects fit the group well yet may fail to adequately describe any individual in the sample, thereby formally attempting to reconcile nomothetic and idiographic perspectives. GIMME has been extensively validated for use in functional neuroimaging and ecological momentary assessment studies (37–39). Here, we adapt a modern variant of GIMME (40) to understand inertia and dynamic transitions among facial EEs using data collected at a high sampling rate, capitalizing on GIMME's ability to handle data collected at high temporal resolutions.

While current technology allows for the collection of facial EE data at extremely high temporal resolutions (30 Hz in our case), our study primarily centers around the dynamic relationships between facial macro-expressions of emotion, which occur roughly 500 ms to 4 seconds after a stimulus (41–43). These macro-expressions differ from micro-expressions—fleeting emotional displays that last about 100 to 500 ms—usually appearing when emotions are intentionally suppressed or hidden (44, 45). Our objective was to leverage the high temporal resolution of our data, with a focus on rapid emotion transitions that are displayed for others to see. Thus, for our purposes, we focus on transitions in facial emotion expressions occurring every 500 ms.

Psychotic disorders, such as schizophrenia, present an opportunity to evaluate the translational potential of our computational approach. Symptoms of these disorders include inexpressivity (e.g., blunted facial expressions) and disorganization (e.g., incongruent affect), which substantially impair functioning and are a major research priority (46–48). These symptoms are often prominent and detected by clinicians through behavioral observation during interview. However, clinician ratings are general and qualitative. More precise analyses of altered EE have relied on labor-intensive methods [e.g., FACS ratings made by human raters (49)], which show associations with clinician-rated symptoms (10, 18, 19). For example, in people with schizophrenia, human ratings of facial expressions are correlated with clinician-rated negative symptoms such as inexpressivity (50, 51), while atypical facial movements are correlated with clinician-rated disorganized symptoms (e.g., inappropriate affect, agitation) (49). Similarly, EMG studies have found that smiling is negatively correlated with clinician-rated negative symptoms (52, 53). In parallel, preliminary applications of FERAs suggest that individuals diagnosed with schizophrenia make fewer pleasant expressions and have less variability in pleasant expressions compared to never-psychotic individuals, and these differences are correlated with clinician-rated negative and disorganized symptoms (30, 54). However, none of these studies have examined EE dynamics.

Overall, this study sought to demonstrate the value of an AI-informed approach to quantifying EE dynamics that can inform our understanding of psychotic disorders and other medical conditions where identifying atypical EE patterns may advance assessment and treatment. Video recordings were collected during a semi-structured clinical interview in an epidemiolocal sample of individuals diagnosed with psychotic disorders and a never-psychotic comparison group (55). EE signals measuring inertia and dynamics were related to gold-standard clinical measures including diagnosis and symptom severity.

Results

We used the FaceReader facial emotion recognition algorithm (26) (see *Methods* for details), which provided normalized intensity ratings (based on the probability of an emotion being present; see *Methods*) of seven EEs—angry, disgusted, happy, neutral, sad, scared, and surprised—over the course of a semi-structured clinical interview (see *Methods* for details). Descriptive statistics for these variables (mean and SD of emotion amplitude), including differences among clinical groups, are provided in *SI Appendix*, Table S1. We fit dynamic network models (CS-GIMME model; details in *Methods*) to the FaceReader-derived EE time series data from each clinical interview, generating a set of autocorrelation parameters (capturing EE inertia), as well as cross-emotion directed associations that reflect lag-1 changes from one EE to another (see Fig. 1 for a conceptual overview of the analytic pipeline). These models yielded EE network structures that characterize common dynamics in the entire sample as well as diagnostic group-specific structures [never-psychotic (NP); other psychosis (OP); and schizophrenia (SZ)].

Network modeling revealed both contemporaneous and lagged relationships between EEs. Lagged effects represent the directed influence of one EE on another over a brief period (here, 500 ms), while contemporaneous effects represent instantaneous associations among EEs (*Methods*). Given our interest in the interplay of EE signals over time, we report only the lagged (dynamic) paths from the network models; contemporaneous effects are provided in *SI Appendix*, Table S2. Group and subgroup networks are visualized in Fig. 2. Individual differences in the strength of specific cross-emotion paths reflect differences in the tendency of subsequent frames to shift from one EE to another. For example, a weaker effect from Neutral to Sad in a given individual would indicate a lower propensity for shifting from neutral toward sad expression at the next time point (Table 1).

Clinical Group Differences in Emotion Expression Inertia. In the full sample, each facial EE had a positive AR_1 coefficient, reflecting that all EEs showed a degree of inertia on average. We tested for group differences in AR_1 coefficients and found that the NP group had a lower Angry AR_1 coefficient compared to the OP group ($d = 0.49$, $P = 0.009$). All other group comparisons in AR_1 coefficients were nonsignificant (Table 2).

Clinical Group Differences in Inter-Emotion Expression Dynamics. Turning to directed effects among EEs, we found significant positive cross-lagged paths from Neutral to Happy, Sad, and

Fig. 1. Analysis pipeline: automated facial emotion expression dynamics. Visual depiction of key steps in an analytic pipeline for automated facial EE dynamics analysis. 1) All participants completed a clinical interview while being recorded for facial movements. 2) Raw video data from the clinical interview was used in conjunction with FaceReader's pre-trained AI model to estimate time-varying facial EEs (30 Hz native resolution). The line plot displays raw data from one representative participant with good data quality. Smoothed bolded lines in the foreground are estimated with a GAM filter for visualization purposes with raw facial emotion signals underlain. 3) Raw EE time series data were preprocessed including interpolation over segments of unusable data and downsampling to a 500ms time grid. Participants with poor data quality were excluded from further analysis (less than 5 min total or 90% or more missing; see *Methods*) 4) Preprocessed data were filtered/pre-whitened with a modified ARMA model prior to dynamic network modeling in order to ensure normally distributed iid residuals after fitting to GIMME. Smoothed raw data (red) are contrasted to full pre-whitening (green) and pre-whitening with refit ARMA coefficients while retaining only lag 1 autoregressive component (blue). 5) Preprocessed and ARMA-filtered time series (retaining AR1 component) were fit to directed network models using CS-GIMME. 6) Path values (partial betas) from the fitted model are extracted from the GIMME model and examined for individual differences based on categorical diagnostic status and continuous psychotic symptom dimensions.

Disgusted that were present across groups (black paths in Fig. 2). Hence, EE patterns showed evidence of moving away from the neutral (primary) state across clinical groups. Furthermore, the network models revealed several instances of EE dynamics that were specific to a clinical group. Individuals in the NP group showed evidence of a tendency to move from Disgusted to Happy, Sad to Happy, and Surprised to Neutral. In the SZ group, a cross-lagged path from Disgusted to Happy was also identified. In addition Neutral to Surprised, Neutral to Scared, and Scared to Sad dynamics were identified in the SZ group. In the OP group, transitions were observed from Happy to Sad, Surprised to Sad, and Disgusted to Sad. Additionally, the OP group showed a negative cross-lagged effect from Sad to Happy, indicating a reduced tendency of transitioning from Sad to Happy on subsequent frames. Magnitudes of these effects are given in Table 3. Altogether, of the 84 cross-lagged effects possible, three

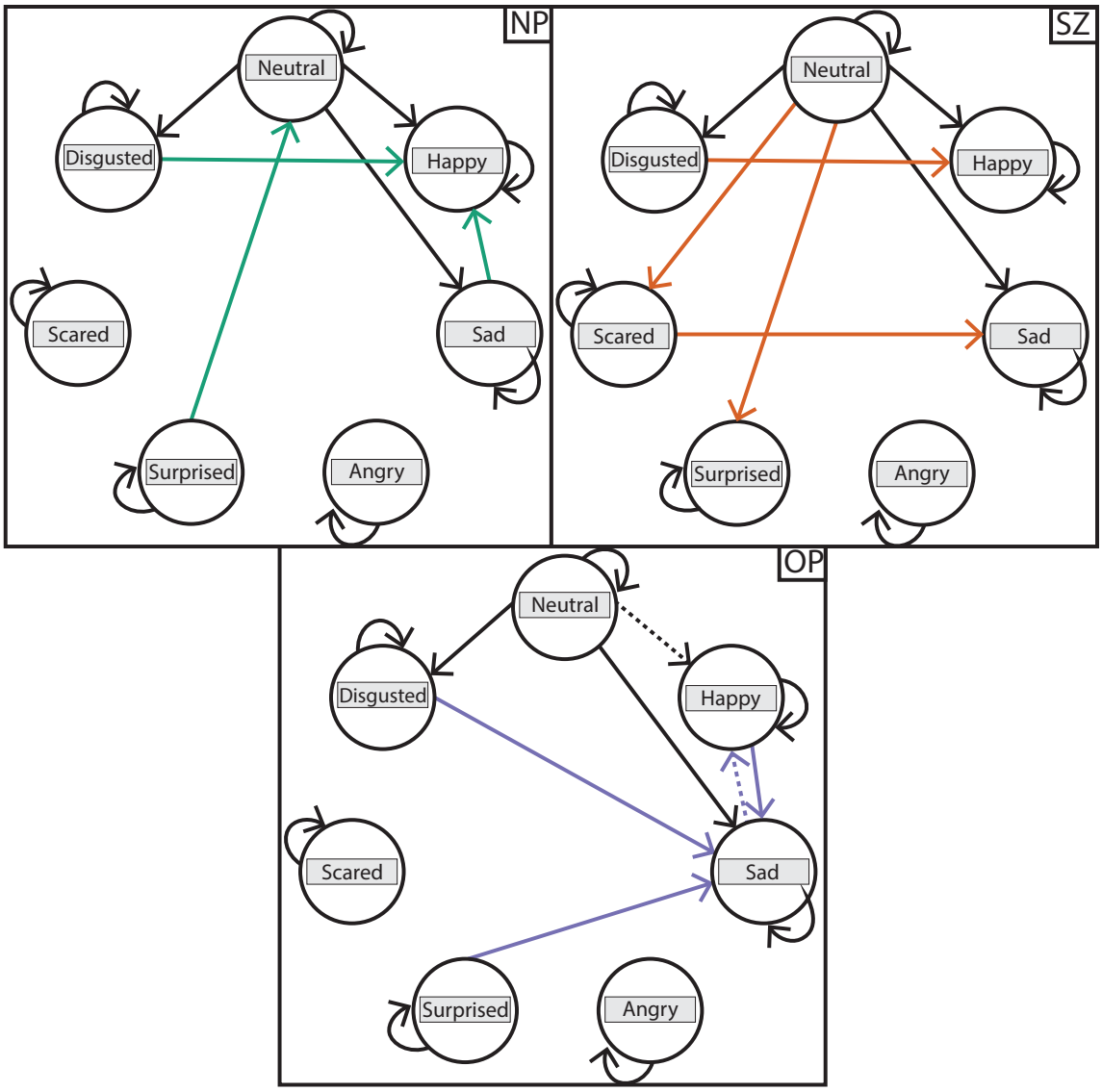

Fig. 2. CS-GIMME network modeling results. Network visualization of cross-lagged CS-GIMME results (contemporaneous paths are omitted, see *Methods*). Black lines represent group-level paths that are estimated for all participants regardless of clinical group assignment, with colored lines denoting subgroup-specific paths that did not improve model fit at the group level but improved fit locally within a clinical group. Dashed lines in the lower (OP) panel denote negative paths, with all others (solid) reflecting positive path estimates. *SZ = Schizophrenia Group, OP = Other Psychosis Group, NP = Never Psychotic Group.*

were observed in all groups, two in two groups, seven in one group, and 72 did not substantially improve model fit and were excluded.

Although these five significant paths (transitions/dynamics) were observed in multiple groups, these effects could vary in magnitude. We tested for such group differences in a mixed-effects regression model (Table 3). We found that Neutral to Happy dynamics were weaker in both SZ and OP groups, but the deficit was particularly pronounced in OP (d = −1.95 and −1.41 compared to NP and SZ, respectively). Neutral to Sad dynamics were greater in OP compared to NP (d = 1.28) and SZ (d = 1.18) groups. Also, transitions from Sad to Happy were less likely in OP compared to NP group (d = −2.27). In contrast, there were no group differences in transition coefficients from Neutral to Disgusted and Disgusted to Happy.

Associations of Emotion Inertia and Dynamics with Symptom Dimensions. We examined associations of EE inertia and dynamics with relevant symptom dimensions by calculating partial correlations that control for study group, sex, age, and race (Fig. 3). We found significant negative partial correlations between disorganization and AR_1 components for Neutral, Happy, Surprised, and Scared, indicating that high disorganization was associated with lower inertia for these expressions. These results indicate that participants with more severe disorganized symptoms had an increased tendency to "leave" these EEs and shift nonspecifically toward other expressions. Further, tardive dyskinesia was associated with weaker Neutral to Sad dynamics, as well as lower Angry and Disgusted inertia. Also, reality distortion was associated with higher Sad AR_1 (inertia).

Finally, we were interested in whether SZ-specific facial emotion dynamics could be attributed to specific psychotic symptoms. In a post hoc exploratory analysis within the SZ group, disorganization was associated with stronger cross-lagged dynamics from Neutral to Scared and Surprised (Table 4). Further, inexpressivity was associated with decreased Neutral to Surprised dynamics and increased Disgusted to Happy dynamics.

Discussion

We developed a computational method that combines FERAs and network modeling to study EE dynamics in video recordings of

Table 1. Baseline demographic and clinical characteristics of final sample

Characteristic	SZ (n = 43) N (%)	OP (n = 53) N (%)	NP (n = 116) N (%)	Case − NP OR (*P*-value)	SZ – OP OR (*P*-value)
Male	25 (58.1)	30 (56.6)	59 (50.9)	0.77	0.94
White	32 (74.4)	45 (84.9)	107 (92.2)	0.43	0.36
Black	3 (7.0)	2 (4.8)	5 (4.3)	1.28	1.85
Other race	8 (18.6)	6 (11.3)	4 (3.4)	3.42	3.13
On antipsychotic	33 (79)	22 (42)	1 (<1)	154.00	5.17
On antidepressant	13 (30)	23 (43)	25 (22)	2.15	0.59
On mood stabilizer	13 (30)	15 (28)	1 (<1)	46.81	1.14
	Mean (SD)	Mean (SD)	Mean (SD)	d (*P*-value)	d (*P*-value)
Age, y	52.0 (8.2)	52.5 (9.3)	56.4 (9.1)	−0.46	−0.06
Reality distortion	5.7 (8.2)	1.0 (2.1)	0.23 (1.1)	0.69	0.82
Disorganization	6.5 (7.6)	3.6 (5.3)	1.6 (2.9)	0.69	0.47
Avolition	17.2 (7.8)	9.5 (7.1)	4.5 (5.6)	1.23	1.05
Inexpressivity	10.2 (9.5)	4.0 (5.5)	1.7 (2.9)	0.88	0.83
Tardive dyskinesia	1.1 (1.6)	1.0 (2.3)	0.1 (0.6)	0.63	0.08

Note. In the upper section, number of participants in each clinical group by demographic characteristics. Medication status was obtained by interview of participants and from their pill bottles. For continuous variables, we report the mean and SD within each clinical group. Reality Distortion is a composite of hallucinations and delusions; Disorganization of formal thought disorder and bizarre behavior; Avolition of apathy, asociality, and engagement in activities; and Inexpressivity of blunted facial and vocal expression and expressive gestures. OR = odds ratio, SZ = Schizophrenia Group, OP = Other Psychosis Group, NP = Never Psychotic Group.

clinical interviews. Our method revealed patterns of EE dynamics that were shared by participants diagnosed with schizophrenia, other psychotic disorders, and never-psychotic adults as well as patterns that differed between these groups. In the other psychosis group (e.g., mood disorders with psychosis), many dynamics converged on Sad, with various EEs moving toward Sad on subsequent timesteps. This was complemented by a reduced tendency to transition from Sad to Happy. This is consistent with prior literature showing that in mood disorders sadness dominates affective expressions (56, 57). Among participants diagnosed with schizophrenia, we found evidence of more frequent facial movements from Neutral to Scared and Surprised. These effects were limited to schizophrenia likely because Scared and Surprised were uncommon expressions more generally, and hence shifts involving these emotions were atypical overall. Importantly, mean levels of scared, surprised, and neutral expressions did not differ between diagnostic groups, but modeling the interplay among EEs was able to elucidate group differences.

In the context of a clinical interview, the observed pattern of transitioning from Neutral to Scared and Surprised might denote several possible explanations. From the perspective that EEs are primarily tools of social influence (17), fearful EEs signal submission to a social partner, which may represent conscious responses to perceived interpersonal threat over the interview. Alternatively, it is possible that the contents of the interview encouraged the discussion of previous frightening experiences. However, we think that our findings primarily reflect a propensity to express uncommon or unexpected EE dynamics. This interpretation is consistent with the literature on the tendency of people with schizophrenia to experience and express emotions that are incongruent with the situation they are in refs. 50, 51, 58, and 59. This explanation was further supported by positive associations between these dynamics and clinician-rated disorganization symptoms (e.g., incongruent affect, bizarre behavior).

Thus, transitions to Surprised and Scared in the SZ group may be primarily indicative of disorganized dynamics. At unusually high levels, stable EE dynamics that tilt the (facial-muscular) system toward these expressions may tap into formal thought disorder measured by SAPS, which may be caused by a breakdown in goal-directed socio-emotional control. Conversely, the pattern of

Table 2. CS-GIMME autoregressive (inertia) coefficients

Emotion	NP Mean (SD)	SZ Mean (SD)	OP Mean (SD)	SZ–NP d	OP–NP d	SZ–OP d
$Neutral_{AR1}$	0.78 (0.10)	0.75 (0.13)	0.75 (0.10)	−0.23	−0.26	0.00
$Happy_{AR1}$	0.72 (0.12)	0.75 (0.14)	0.75 (0.23)	0.17	0.18	−0.03
Sad_{AR1}	0.70 (0.15)	0.64 (0.19)	0.68 (0.14)	−0.40†	−0.19	−0.21
$Angry_{AR1}$	0.68 (0.20)	0.71 (0.23)	**0.77 (0.16)**	0.12	**0.49****	−0.35†
$Surprised_{AR1}$	0.65 (0.19)	0.68 (0.18)	0.71 (0.15)	0.15	0.35	−0.21
$Scared_{AR1}$	0.64 (0.19)	0.64 (0.20)	0.63 (0.16)	0.00	−0.06	0.06
$Disgusted_{AR1}$	0.72 (0.13)	0.72 (0.15)	0.72 (0.15)	0.03	0.01	−0.03

Note. Mean and SD values for each inertia (AR_1) coefficient calculated for each clinical group. Righthand columns show the effect size of and group difference comparison between pairs of subgroups, expressed in Cohen's d. Statistical significance (FDR-corrected) is denoted: $^{***}P \le 0.005$, $^{**}P \le 0.01$, $^{*}P \le 0.05$, $^{\dagger}P \le 0.10$. SZ = Schizophrenia Group, OP = Other Psychosis Group, NP = Never Psychotic Group.

Table 3. CS-GIMME cross-lagged (inter-EE dynamics) coefficients

		NP	SZ	OP	SZ-NP	OP-NP	OP-SZ
Edge	Group	Mean (SD)	Mean (SD)	Mean (SD)	d	d	d
Neutral → Disgusted	All	0.29 (0.17)	0.27 (0.19)	0.27 (0.18)	−0.14	−0.13	−0.01
Neutral → Happy	**All**	**0.30 (0.18)**	**0.23 (0.20)**	**−0.01 (0.14)**	**−0.38****	**−1.95*****	**−1.41*****
Neutral → Sad	**All**	**0.22 (0.16)**	**0.20 (0.21)**	**0.45 (0.24)**	−0.16	**1.28*****	**1.18*****
Sad → Happy	**NP, OP**	**0.13 (0.14)**	**0**	**−0.37 (0.40)**	**–**	**−2.27*****	**–**
Disgusted→ Happy	NP, SZ	0.15 (0.14)	0.14 (0.15)	0	−0.04	–	–
Surprised→ Neutral	NP	0.10 (0.09)	0	0	–	–	–
Neutral → Surprised	SZ	0	0.09 (0.12)	0	–	–	–
Neutral → Scared	SZ	0	0.13 (0.15)	0	–	–	–
Scared→ Sad	SZ	0	0.07 (0.09)	0	–	–	–
Happy → Sad	OP	0	0	0.48 (0.38)	–	–	–
Surprised→ Sad	OP	0	0	0.13 (0.19)	–	–	–
Disgusted→ Sad	OP	0	0	0.25 (0.21)	–	–	–

Note. Mean and SD values for each cross-lagged coefficient calculated within each clinical group. Group column represents what subgroup the transition coefficient was estimated for (With "All" denoting coefficients that were estimated for all groups). Pairwise group differences are expressed in Cohen's d. Statistical significance (FDR-adjusted) is denoted: $^{***}P \le 0.005$, $^{**}P \le 0.01$, $^{*}P \le 0.05$, $^{\dagger}P \le 0.10$. SZ = Schizophrenia Group, OP = Other Psychosis Group, NP = Never Psychotic Group.

transitions to Sad in the OP group was not related to a specific dimension of psychosis and is likely more attributable to concurrent (typically primary) mood pathology in this group (e.g., schizoaffective disorder, mood disorder with psychosis).

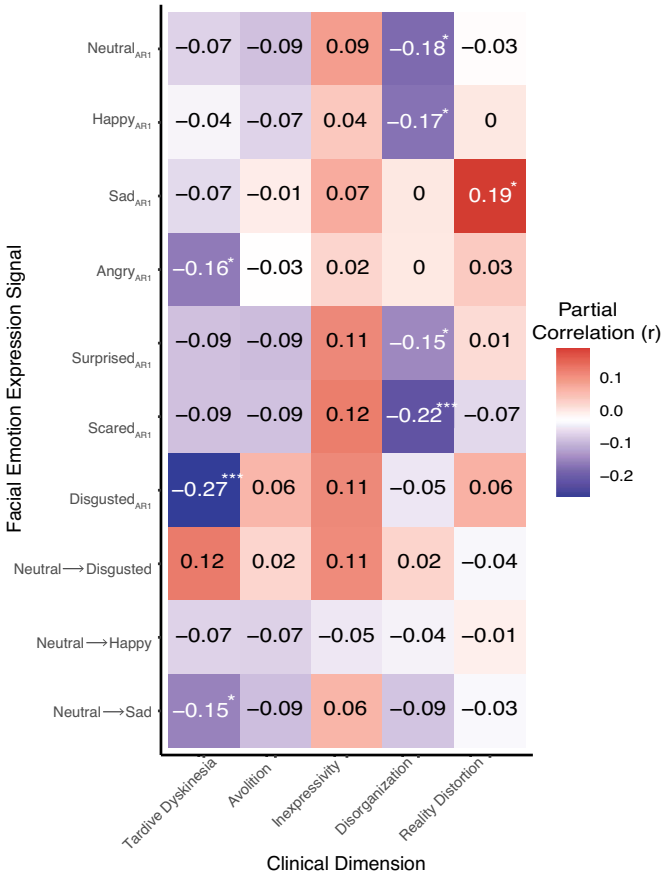

Fig. 3. Results of individual differences analysis. Note. Heatplot of partial correlation coefficients between dimensional psychosis scores and group-level facial EE signals (covariates: clinical group and demographic variables). Significant results are highlighted with white text and asterisks ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.005$). All *P*-values have been FDR adjusted. See *SI Appendix*, Fig. S4 for unconditional (bivariate) correlation heatplot.

Somewhat unexpectedly, we did not find strong group differences in EE inertia. However, inertia was related to specific symptom dimensions. Most notably, greater disorganization was associated with lower inertia in Neutral, Happy, Surprised, and Scared expressions. This pattern suggests that people with more volatile facial expressions were rated by interviewers as more disorganized. This is consistent with findings of ecological momentary assessment research that disorganization is linked to greater variability in negative affect over time (60). Interestingly, clinician-rated inexpressivity was weakly associated with greater inertia of each EE, indicating reduced variability of overall facial expression, although this effect did not reach statistical significance. Tardive dyskinesia was associated with multiple abnormalities in expression, most significant of which was volatility (low inertia) in Disgust. Indeed, tardive dyskinesia often affects muscles of facial expression (61), suggesting that these particular EE dynamics may have neurological origins at least in part. This underscores utility of our computational approach beyond psychiatry, as it has potential applications in screening for certain neurologic conditions or side effects of prolonged antipsychotic use.

Methodologically, the conjunction of FERAs with dynamic network models is the natural confluence of two mature computational approaches that can be harmonized for use in research and applied settings. This technological confluence is timely given availability of video data resulting from the rise of telemedicine during the COVID-19 pandemic (62). While large amounts of video data are easily collectible or already in existence, a key missing element has been a scalable and straightforward analytic framework to extract the useful EE content contained within this information-rich data modality. We described such a framework and demonstrated that the EE dynamics it identified relate to psychiatric diagnosis and symptom measures. Further, our method captures fine-grained (~500 ms) EE dynamics that would be extremely labor-intensive to collect using human raters. This opens the possibility for application to different contexts [e.g., brief video collection in daily life (63, 64), disorders (affective and personality disorders), and clinical questions [assistance with diagnosis, treatment-outcome measurement, guidance of just-in-time adaptive interventions (65)].

These strengths notwithstanding, we note limitations of this work that can guide future efforts to validate and expand upon

Table 4. Post-hoc analysis: Associations between SZ-specific facial EE signals and psychosis dimensions.

Facial EE signal	Clinical dimension	Estimate	S.E.	*t*	*P*
Disgusted→ Happy	**Inexpressivity**	**18.47**	**7.43**	**2.49**	**0.01**
(SZ-only)	Avolition	1.83	7.44	0.25	0.81
	Reality distortion	−3.13	7.45	−0.42	0.67
	Disorganization	−8.11	7.43	−1.09	0.28
	Tardive dyskinesia	0.37	7.46	0.05	0.96
Scared→ Sad	Inexpressivity	17.91	12.85	1.40	0.16
(SZ-only)	Avolition	−21.76	12.87	−1.69	0.09
	Reality distortion	−2.53	12.89	−0.20	0.84
	Disorganization	−19.44	12.85	−1.51	0.13
	Tardive dyskinesia	−2.62	13.13	−0.20	0.84
Neutral→ Scared	Inexpressivity	12.45	7.26	1.71	0.09
(SZ-only)	Avolition	0.81	7.26	0.11	0.91
	Reality distortion	13.93	7.50	1.86	0.06
	Disorganization	**15.15**	**7.26**	**2.09**	**0.04**
	Tardive dyskinesia	1.35	7.41	0.18	0.86
Neutral→ Surprised	**Inexpressivity**	**−19.04**	**9.05**	**−2.10**	**0.04**
(SZ-only)	Avolition	−0.52	9.11	−0.06	0.95
	Reality distortion	21.15	14.49	1.46	0.14
	Disorganization	**21.93**	**9.05**	**2.42**	**0.02**
	Tardive dyskinesia	−3.55	9.10	−0.39	0.70

Note. Regression coefficients, with SE, t-statistics, and *P*-values for post hoc analyses relating SZ-only edges to dimensions of psychosis. All *p*-values are FDR-corrected. SZ = Schizophrenia Group, OP = Other Psychosis Group, NP = Never Psychotic Group.

our method. First, inferences about the dynamics of emotional experience are limited by our method's focus on facial expressions. This can be addressed in part by extending the framework to consider the content of what the participant said. Linguistic characteristics (e.g., arousal, valence, topic) of the interviewee's responses can be included in GIMME models alongside facial expressions. AI models are already available that assess these characteristics rather accurately (66), though incorporating linguistic content alongside EEs will require modifications to the present framework and further validation. Second, the sample was largely White. Although there is evidence that EEs generalize across racial groups (22), a recent study found that FaceReader may fail to classify facial expressions more frequently for Black than non-Black participants (67). Though we did not find this pattern in our data (*SI Appendix, Materials*), the present findings need to be replicated in more diverse samples. Finally, our analysis was limited to emotion dynamics that occurred at fast timescales, and slower dynamics that manifest seconds to minutes later would be missed in our approach. Despite these limitations, we provide proof-of-concept for the general computational approach and demonstrate the ability of this method to explain variation in relevant dimensions of psychosis.

Facial expressions are fundamentally intertwined with emotional experience and the desire to convey or hide an internal experience in a social setting. However, until recently, the ability to quantify facial expressions at the scale needed to examine EE dynamics has remained a challenge. We presented a computational method for quantifying dynamics among EEs that addresses this bottleneck. With medicine's increasing utilization of telehealth and AI, FERAs and complementary computational methods are poised to serve an increasing role in quantifying objective indices of socio-emotional health. For example, automated EE detection may provide actionable read-outs of emotional experience or social signaling that could help clinicians identify the onset or worsening of psychopathology (68), which could indicate a need for referrals to a psychiatrist or neurologist or the need for pain management. Likewise, facial emotion profiles linked to diagnoses, symptoms, or other clinical features can help in treatment selection by augmenting data collected from clinical interviews with signals that may be too subtle for human practitioners to notice in the aggregate. Further refinement of this approach may prove valuable in identifying objective markers of socio-emotional abnormalities that can inform clinical judgment and treatment in psychology, psychiatry, and medicine more broadly.

Methods

Sample. Participants were drawn from the Suffolk County Mental Health Project, a longitudinal study of first-admission psychosis recruited 1989–1995 from the 12 inpatient units in Suffolk County, NY (response rate 72%; Bromet, 1992). Eligibility criteria included residence in Suffolk County, age between 15 and 60, ability to speak English, IQ > 70, first admission within the past 6 mo, current psychosis, and no apparent medical etiology for psychotic symptoms. Research diagnoses were made by consensus of study psychiatrists using all available information collected by Year 20, including medical records, significant other interviews, and diagnostic interviews. The diagnostic process has been described previously (69). For the present study, groups of interest were SZ and other psychosis (OP). A sex- and age-matched comparison group of 261 NP adults was recruited using random digit dialing in zip codes where members of the psychosis cohort resided at Year 20 (response rate 67%) (55, 70). Written informed consent was obtained at each visit. This research was approved annually by the Committee on Research Involving Human Subjects at Stony Brook University. Participants received financial compensation for their time.

The 25-y follow-up included 569 participants. The current study is based on participants who consented to video recording and the recording had sufficient quality for FERA + GIMME analysis (see below). This resulted in 43 SZ, 53 OP, and 116 NP participants. Table 1 describes the demographic characteristics of the three groups.

Procedure. Video-recordings were taken during a clinical interview that consisted of the Structured Clinical Interview for DSM-IV (71) and Quality of Life Scale interview (72).

Clinical Measures. Symptoms were assessed using the Schedule for the Assessment of Positive Symptoms (SAPS) (73) and Schedule for the Assessment of Negative Symptoms (SANS) (74). Interviewers rated symptoms based on all available information, including data collected during the clinical interview, information from medical records, and interviews with significant others who know the participant well. The SAPS and SANS were scored into four factor-analytically derived subscales–reality distortion, disorganization, inexpressivity, and avolition–described previously (46). Internal consistency of these scales was in the acceptable range ($\alpha = 0.77$ to 0.88). Tardive dyskinesia was assessed as a total score on the Abnormal Involuntary Movement Scale (75), which includes seven observational ratings of movements in face, lips, tongue, extremities, and trunk.

Facial Emotion Time Series Generation and Preprocessing. Participant videos were translated into facial EE time series using FaceReader version 8.0 (26): a commercially available automated facial emotion classification software. FaceReader is recognized as one of the most accurate FERAs with comparison studies demonstrating excellent convergence with manually coded facial affect (31), and is 79% as accurate as human raters even under naturalistic conditions (16). FaceReader uses a deep neural network (76, 77) trained on a set of over 20,000 manually annotated images to classify continuously measured facial displays across seven basic emotions identified by Paul Ekman (78, 79). This training dataset was racially diverse, and FaceReader showed excellent performance in two independent samples that included Asian, Black, and Hispanic participants (16). FaceReader generates an estimate of the degree of concordance of the current facial expression with the neural network's learned/cached representation of each EE, returning the probability of each EE from 0 to 1*. Time series for each of the seven EEs are generated at a sampling rate of 30 Hz.

Prior to analysis, we combined the raw FaceReader output for participants with more than one video file into a single data frame. Given our interest in primarily characterizing AR1 processes in these data, we concatenated all data frames and padded transitions between video files with 1,000 missing values (so as not to affect GIMME). We smoothed over very brief periods of unusable data with Stineman interpolation (80), implemented in the *imputeTS* R package (81). We interpolated over a conservative maximum window of 30 frames/1 s of missing data to ensure that interpolated values were anchored by reliable EE estimates. After interpolating over small segments of missing data, we excluded participants that either had five minutes or less *or* contained less than 10% of readable frames. Distributions of total valid frames and of the overall percentage of valid frames are presented in *SI Appendix*, Table S2 and Fig. S3. Out of an initial 240 participants, we excluded 28 ($n_{SZ} = 12$, $n_{OP} = 7$, $n_{NP} = 9$) based on these criteria, leaving 212 participants for CS-GIMME fitting. Given our a priori interest in studying rapidly unfolding EE dynamics that are interpersonally apparent (i.e. macroexpressions), we downsampled from 30 Hz to 2 Hz. Prior to downsampling, we observed relatively smooth transitions in autocorrelation and cross-correlations functions (ACF, CCF, respectively; see *SI Appendix*, Fig. S2) across lags, providing evidence that cross-lagged emotional transitions did not preferentially dominate any specific frequency band across EEs. Downsampling was performed by chunking 30 Hz FaceReader output into 15-frame bins and taking the mean of non-missing values for every EE. Downsampled time series were then within-participant standardized to a mean of 0 and SD of 1.

Preparation for network modeling. Prior to fitting CS-GIMME, we prewhitened each EE timeseries using an Auto-Regressive Moving Average (ARMA[8,2]; see *SI Appendix* for details) model to avoid contamination of effects of interest (lag-1) with autoregressive effects from higher-order lagged influences (lag-2, lag-3, etc.). However, given that CS-GIMME explicitly models AR_1 components of the multivariate timeseries, we extracted the ARMA coefficients for all EEs and refit the downsampled timeseries, this time with the AR_1 coefficient fixed to 0 and retained the residuals of this model to be fit by CS-GIMME. This procedure effectively removes higher-order autoregressive properties of the downsampled timeseries, while preserving the lag-1 autocorrelation structure of the data (figure. 2, 4), which is explicitly modeled within the CS-GIMME framework. This final preprocessing step ensures that GIMME residuals are *iid* and normally distributed and conform to model assumptions.

Network Model Fitting. Preprocessed and downsampled EE timeseries were used to estimate contemporaneous (lag-0) and lagged relationships among facial expressions using the Confirmatory Subgrouping Group Iterative Multiple Model Estimation algorithm [CS-GIMME (37, 40)]. The GIMME algorithm identifies individual and group connectivity patterns by fitting structural equation models† to time series data, and iteratively adjusting models for improved fit using modification indices in two steps: a whole group-level search followed by an individual-level search. CS-GIMME extends the GIMME algorithm by performing an additional subgroup-specific search for paths between timeseries that are present in a labeled subgroup, but not in others. Confirmatory subgroups were SZ, OP, and NP. Paths estimated for the entire sample were estimated if their addition into the structural model improved model fit for ≥75% of participants in the total sample (39, 82). Subgroup-specific paths were added if they improved fit for ≥65% participants in the group (this cutoff was determined empirically; see *SI Appendix, Methods* and Fig. S5).

Analytic Approach.

Group differences in facial emotion inertia and dynamics. After fitting CS-GIMME to EE timeseries, we extracted estimated regression coefficients for cross-lagged effects and treated these as edges in participant-specific graphs. Thus, participant-specific network structures consisted of a combination of edges that were estimates for all participants (whole group-level), participants within an a priori subgroup (subgroup-level), and individual-level edges. As noted above, nodes in these networks are each of Eckman's facial EE types and can generally be thought of as "EE states", though note that facial emotion signals are computed as continuous measures and do not undergo state transitions in a Markovian sense. In this continuous representation, it is plausible to conceptualize EE signals as the intensity of a particular EE. Thus, the extracted betas from CS-GIMME more accurately represent the parametric weight on a participant's tendency to transition in intensity from one EE to another (see *SI Appendix, Materials* for an in-depth discussion on how our approach differs from discrete state-space approaches such as hidden Markov models).

CS-GIMME estimates both contemporaneous (lag-0) and cross-lagged (lag-1) graph edges. The contemporaneous edges represent the relationships among EEs over the previous 0.5 s, which tend to be suppressive (i.e., high neutral scores are associated with lower happy, scared, angry, etc scores within the same timepoint). We were primarily interested in group differences in the lagged networks, given that the primary statistical innovation in using CS-GIMME (and dynamic network models more generally) lies in its ability to analyze inter- and intra-EE dynamics from time t to time $t + 1$. Due to our primary interest in temporal dynamics (inertia and transitions), we conceptualized contemporaneous associations as absorbing nuisance variation. As such, under the current analysis regime, contemporaneous associations partial out nuisance variation at the same timepoint to promote EE specificity in lag-1 transitions. In other words, lagged effects reflect the *unique* tendency to transition from one EE to another.

In the current analysis, we were interested in two primary measures of the lagged network: the AR_1 component of each EE time series and weights for cross-lagged edges estimated at the group and subgroup level. First, we interpret higher AR_1 paths for a given EE as greater inertia of the expression or a heightened propensity to remain relatively stable with respect to the expression of an emotion over the course of the interview (34). We additionally analyzed group differences in common cross-lagged edges (betas) as a test of whether transitions among specific EEs differed between groups.

Primary data analysis consisted of two similarly structured mixed effects regression models, where graph measures (AR_1 components, cross-lagged edge values) were predicted by interacting categorical predictors of clinical group (NP, OP, SZ) and the specific EE dynamics parameters (in the AR_1 model this was a categorical predictor denoting whether the value was an angry, sad, neutral, etc. AR_1 path). We additionally included measures of video quality (n valid frames, n complete frames, percentage of missing frames) as nuisance

*Of note, the goal of our analysis is to quantify smooth transitions amongst EEs in a continuous representation (rather than discretizing frames into specific expression "states"), which maximizes statistical power and precision. As we note below, in a continuous setting we can consider these measures of expression intensity.

†More specifically, GIMME uses the extended unified structural equation model [euSEM (86)], where contemporaneous and lagged paths are estimated simultaneously.

covariates. Pairwise contrasts between groups were tested, controlling for the number of comparisons (83, 84). Degrees of freedom for contrasts were computed using the Kenward–Roger method, and *p*-values were adjusted using the Tukey method for comparing each family of three tests (NP-SZ, NP-OP, SZ-OP).

Associations with Variation in Psychosis Symptom Dimensions. After examining group differences in CS-GIMME parameters (AR_1 paths and directed edge values), we tested whether these variables were related to symptom dimensions to investigate more specific clinical constructs than heterogeneous disorders. Schizophrenia and psychosis-spectrum disorders are heterogenous in their symptom presentation, meaning that the presence or absence of a group-level effect does not preclude more focal associations between a measure of facial affect dynamics and a specific domain of psychotic functioning. We computed the partial Pearson correlation of each common CS-GIMME attribute with the five clinical measures, controlling for clinical group status, biological sex, and race. Statistical control for study group ensured that the partial correlations are not redundant with effects of group status, but rather reflect incremental dimensional associations beyond diagnostic group.

Post-hoc Subgroup-Specific Analyses. Finally, we ran a series of post hoc mixed effects models based on subgroup-specific paths estimated by CS-GIMME. While subgroup edges denote that most participants in a given a priori subgroup display evidence of significant EE dynamics, we sought to directly test whether these subgroup-specific effects showed a cleaner mapping onto specific symptom dimensions. With respect to SZ-specific edges, we ran a separate mixed effects model predicting clinical dimension scores by edge values. Given that overall, these edges are absent among non-SZ participants, we ran these models using a filtered dataset containing only SZ participants. Significant associations in these models indicate that *within participants with SZ*, higher subgroup-specific edges denote greater psychotic symptom severity, serving as a post hoc "dimensional specificity analysis."

Data, Materials, and Software Availability. Anonymized between-subjects variables used in our analyses, along with computed facial emotion signals can be found on the Open Science Framework repository (https://osf.io/8gsye/) (85). An R package with usable functions extracted from our analysis is publicly available on GitHub, which is linked in the OSF repository.

ACKNOWLEDGMENTS. This work was supported by the NIH (R01MH110434 to R.K.). We thank Evelyn Bromet, founder of the cohort. We gratefully acknowledge the support of the participants and mental health community of Suffolk County for contributing their time and energy to this project. We are also indebted to dedicated efforts of study coordinators, interviewers for their careful assessments, and to the psychiatrists who derived the consensus diagnoses.

1. B. F. Gruenbaum *et al.*, Absence seizures and their relationship to depression and anxiety: Evidence for bidirectionality. *Epilepsia* **62**, 1041–1056 (2021).
2. D. K. Y. Leung, W. C. Chan, A. Spector, G. H. Y. Wong, Prevalence of depression, anxiety, and apathy symptoms across dementia stages: A systematic review and meta-analysis. *Int. J. Geriatr. Psychiatry* **36**, 1330–1344 (2021).
3. M. R. Frumkin, T. L. Rodebaugh, The role of affect in chronic pain: A systematic review of within-person symptom dynamics. *J. Psychosom. Res.* **147**, 110527 (2021).
4. A. M. Kring, "Emotion disturbances as transdiagnostic processes in psychopathology" in *Handbook of Emotions* (The Guilford Press, ed. 3, 2008), pp. 691–705.
5. K. M. Prkachin, Assessing pain by facial expression: Facial expression as nexus. *Pain Res. Manag.* **14**, 53–58 (2009).
6. H. Davies *et al.*, Facial expression to emotional stimuli in non-psychotic disorders: A systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* **64**, 252–271 (2016).
7. J. D. Henry, P. G. Rendell, A. Scicluna, M. Jackson, L. H. Phillips, Emotion experience, expression, and regulation in Alzheimer's disease. *Psychol. Aging* **24**, 252–257 (2009).
8. M. Bologna *et al.*, Facial bradykinesia. *J. Neurol. Neurosurg. Psychiatry* **84**, 681–685 (2013).
9. B. Jin, Y. Qu, L. Zhang, Z. Gao, Diagnosing parkinson disease through facial expression recognition: Video analysis. *J. Med. Internet Res.* **22**, e18697 (2020).
10. C. G. Kohler, E. A. Martin, Emotional processing in schizophrenia. *Cogn. Neuropsych.* **11**, 250–271 (2006).
11. W. T. Carpenter Jr., R. W. Buchanan, Negative symptom therapeutics. *Schizophr. Bull.* **43**, 681–682 (2017).
12. D. Keltner, A. M. Kring, Emotion, social function, and psychopathology. *Rev. Gen. Psychol.* **2**, 320–342 (1998).
13. T. Hadjistavropoulos, K. D. Craig, A theoretical framework for understanding self-report and observational measures of pain: A communications model. *Behav. Res. Ther.* **40**, 551–570 (2002).
14. T. S. Conner, L. F. Barrett, Trends in ambulatory self-report: The role of momentary experience in psychosomatic medicine. *Psychosom. Med.* **74**, 327–337 (2012).
15. T. J. Trull, S. P. Lane, P. Koval, U. W. Ebner-Priemer, Affective dynamics in psychopathology. *Emot. Rev.* **7**, 355–361 (2015).
16. D. Dupré, E. G. Krumhuber, D. Küster, G. J. McKeown, A performance comparison of eight commercially available automatic classifiers for facial affect recognition. *PLoS One* **15**, e0231968 (2020).
17. C. Crivelli, A. J. Fridlund, Facial displays are tools for social influence. *Trends Cogn. Sci.* **22**, 388–399 (2018).
18. A. M. Kring, E. K. Moran, Emotional response deficits in Schizophrenia: Insights from affective science. *Schizophr. Bull.* **34**, 819–834 (2008).
19. M. K. Mandal, U. Habel, R. C. Gur, Facial expression-based indicators of schizophrenia: Evidence from recent research. *Schizophr. Res.* **252**, 335–344 (2023).
20. C. Darwin, *The Expression of the Emotions in Man and Animals* (University of Chicago Press, 1872).
21. C. E. Izard, Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychol. Bull.* **115**, 288–299 (1994).
22. A. S. Cowen *et al.*, Sixteen facial expressions occur in similar contexts worldwide. *Nature* **589**, 251–257 (2021).
23. P. Ekman, W. V. Friesen, *Facial Action Coding System: Manual* (Consulting Psychologists Press, 1978).
24. K. Wolf, Measuring facial expression of emotion. *Dialogues Clin. Neurosci.* **17**, 457–462 (2015).
25. I. R. Galatzer-Levy, J.-P. Onnela, Machine learning and the digital measurement of psychological health. *Annu. Rev. Clin. Psychol.* **19**, 133–154 (2023).
26. Noldus, FaceReader: Tool for Automatic Analysis of Facial Expression (Version 8.0., Noldus Information Technology B.V., Wageningen, the Netherlands, 2014).
27. S. Stöckli, M. Schulte-Mecklenbeck, S. Borer, A. C. Samson, Facial expression analysis with AFFDEX and FACET: A validation study. *Behav. Res. Methods* **50**, 1446–1460 (2018).
28. F. Z. Canal *et al.*, A survey on facial emotion recognition techniques: A state-of-the-art literature review. *Inf. Sci.* **582**, 593–617 (2022).
29. T. Cowan *et al.*, Computerized analysis of facial expressions in serious mental illness. *Schizophr. Res.* **241**, 44–51 (2022).
30. A. S. Cohen *et al.*, Ambulatory digital phenotyping of blunted affect and alogia using objective facial and vocal analysis: Proof of concept. *Schizophr. Res.* **220**, 141–146 (2020).
31. T. Gupta *et al.*, Alterations in facial expressions of emotion: Determining the promise of ultrathin slicing approaches and comparing human and automated coding methods in psychosis risk. *Emotion* **22**, 714–724 (2022), 10.1037/emo0000819.
32. P. Werner *et al.*, Automatic recognition methods supporting pain assessment: A survey. *IEEE Trans. Affect. Comput.* **13**, 530–552 (2022).
33. B. Sonawane, P. Sharma, Review of automated emotion-based quantification of facial expression in Parkinson's patients. *Vis. Comput.* **37**, 1151–1167 (2021).
34. P. Koval, P. T. Burnett, Y. Zheng, "Emotional inertia: On the conservation of emotional momentum" in *Affect Dynamics*, C. E. Waugh, P. Kuppens, Eds. (Springer International Publishing, 2021), pp. 63–94, 10.1007/978-3-030-82965-0_4.
35. P. Kuppens, P. Verduyn, Emotion dynamics. *Curr. Opin. Psychol.* **17**, 22–26 (2017).
36. J. J. Gross, Emotion regulation: Current status and future prospects. *Psychol. Inq.* **26**, 1–26 (2015).
37. K. M. Gates, P. C. M. Molenaar, Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage* **63**, 310–319 (2012).
38. S. T. Lane, K. M. Gates, H. K. Pike, A. M. Beltz, A. G. C. Wright, Uncovering general, shared, and unique temporal patterns in ambulatory assessment data. *Psychol. Methods* **24**, 54–69 (2019).
39. A. M. Beltz, K. M. Gates, Network mapping with GIMME. *Multivar. Behav. Res.* **52**, 789–804 (2017).
40. T. R. Henry *et al.*, Comparing directed functional connectivity between groups with confirmatory subgrouping GIMME. *NeuroImage* **188**, 642–653 (2019).
41. P. Ekman, W. V. Friesen, Nonverbal leakage and clues to deception. *Psychiatry* **32**, 88–106 (1969).
42. P. D. P. Ekman, *Emotions Revealed, Second Edition: Recognizing Faces and Feelings to Improve Communication and Emotional Life* (Holt Paperbacks, 2007).
43. U. Hess, R. E. Kleck, Differentiating emotion elicited and deliberate emotional facial expressions. *Eur. J. Soc. Psychol.* **20**, 369–385 (1990).
44. E. Svetieva, Seeing the unseen: Explicit and implicit communication effects of naturally occuring emotion microexpressions (State University of New York at Buffalo, 2014).
45. W.-J. Yan, Q. Wu, J. Liang, Y.-H. Chen, X. Fu, How fast are the leaked facial expressions: The duration of micro-expressions. *J. Nonverbal Behav.* **37**, 217–230 (2013).
46. R. Kotov *et al.*, Validating dimensions of psychosis symptomatology: Neural correlates and 20-year outcomes. *J. Abnorm. Psychol.* **125**, 1103–1119 (2016).
47. E. Roche, L. Creed, D. MacMahon, D. Brennan, M. Clarke, The epidemiology and associated phenomenology of formal thought disorder: A systematic review. *Schizophr. Bull.* **41**, 951–962 (2015).
48. G. P. Strauss, A. S. Cohen, A transdiagnostic review of negative symptom phenomenology and etiology. *Schizophr. Bull.* **43**, 712–719 (2017).
49. C. G. Kohler *et al.*, Static posed and evoked facial expressions of emotions in schizophrenia. *Schizophr. Res.* **105**, 49–60 (2008).
50. G. Bersani *et al.*, Comparison of facial expression in patients with obsessive-compulsive disorder and schizophrenia using the Facial Action Coding System: A preliminary study. *Neuropsychiatr. Dis. Treat.* **8**, 537–547 (2012).
51. C. G. Kohler *et al.*, Dynamic evoked facial expressions of emotions in schizophrenia. *Schizophr. Res.* **105**, 30–39 (2008).
52. K. Wolf, R. Mass, F. Kiefer, K. Wiedemann, D. Naber, Characterization of the facial expression of emotions in Schizophrenia patients: Preliminary findings with a new electromyography method. *Can. J. Psychiatry* **51**, 335–341 (2006).
53. M. Riehle, T. M. Lincoln, Social consequences of subclinical negative symptoms: An EMG study of facial expressions within a social interaction. *J. Behav. Ther. Exp. Psychiatry* **55**, 90–98 (2017).
54. T. Tron, A. Peled, A. Grinsphoon, D. Weinshall, "Automated facial expressions analysis in Schizophrenia: A continuous dynamic approach" in *Pervasive Computing Paradigms for Mental Health*, S. Serino, A. Matic, D. Giakoumis, G. Lopez, P. Cipresso, Eds. (Springer International Publishing, 2016), pp. 72–81, 10.1007/978-3-319-32270-4_8.

55. K. Jonas *et al.*, The course of general cognitive ability in individuals with psychotic disorders. *JAMA Psychiatry* **79**, 659–666 (2022).
56. L. F. Bringmann, L. H. J. M. Lemmens, M. J. H. Huibers, D. Borsboom, F. Tuerlinckx, Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychol. Med.* **45**, 747–757 (2015).
57. K. U. Likowski *et al.*, Sad and lonely? Sad mood suppresses facial mimicry. *J. Nonverbal Behav.* **35**, 101–117 (2011).
58. A. S. Cohen, K. S. Minor, Emotional experience in patients with Schizophrenia revisited: Meta-analysis of laboratory studies. *Schizophr. Bull.* **36**, 143–150 (2010).
59. A. H. Sanchez, L. M. Lavaysse, J. N. Starr, D. E. Gard, Daily life evidence of environment–incongruent emotion in schizophrenia. *Psychiatry Res.* **220**, 89–95 (2014).
60. K. C. Kemp, S. H. Sperry, L. Hernández, N. Barrantes-Vidal, T. R. Kwapil, Affective dynamics in daily life are differentially expressed in positive, negative, and disorganized schizotypy. *J. Psychopathol. Clin. Sci.* **132**, 110–121 (2023).
61. A. M. Anusa *et al.*, A study on drug-induced tardive dyskinesia: Orofacial musculature involvement and patient's awareness. *J. Orofac. Sci.* **10**, 86–95 (2018).
62. C. Cortez, O. Mansour, D. M. Qato, R. S. Stafford, G. C. Alexander, Changes in short-term, long-term, and preventive care delivery in US office-based and telemedicine visits during the COVID-19 pandemic. *JAMA Health Forum* **2**, e211529 (2021).
63. T. Cowan, G. P. Strauss, I. M. Raugh, T. P. Le, A. S. Cohen, How do social factors relate to blunted facial affect in schizophrenia? A digital phenotyping study using ambulatory video recordings *J. Psychiatr. Res.* **150**, 96–104 (2022).
64. A. Abbas *et al.*, Facial and vocal markers of Schizophrenia measured using remote smartphone assessments: Observational study. *JMIR Form. Res.* **6**, e26276 (2022).
65. I. Nahum-Shani *et al.*, Just-in-time adaptive interventiOns (jitais) in mobile health: Key components and design principles for ongoing health behavior support. *Ann. Behav. Med.* **52**, 446–462 (2016), 10.1007/s12160-016-9830-8.
66. J. C. Eichstaedt *et al.*, Closed- and open-vocabulary approaches to text analysis: A review, quantitative comparison, and recommendations. *Psychol. Methods* **26**, 398–427 (2021).
67. T. Cowan, A. S. Cohen, I. M. Raugh, G. P. Strauss, Ambulatory audio and video recording for digital phenotyping in schizophrenia: Adherence & data usability. *Psychiatry Res.* **311**, 114485 (2022).
68. B. Nelson, P. D. McGorry, M. Wichers, J. T. W. Wigman, J. A. Hartmann, Moving from static to dynamic models of the onset of mental disorder: A review. *JAMA Psychiatry* **74**, 528–534 (2017).
69. E. J. Bromet *et al.*, Diagnostic shifts during the decade following first admission for psychosis. *Am. J. Psychiatry* **168**, 1186–1194 (2011).
70. E. Velthorst *et al.*, The 20-year longitudinal trajectories of social functioning in individuals with psychotic disorders. *Am. J. Psychiatry* **174**, 1075–1085 (2017).
71. R. L. Spitzer, J. B. Williams, M. Gibbon, M. B. First, The Structured Clinical Interview for DSM-III-R (SCID). I: History, rationale, and description *Arch. Gen. Psychiatry* 49, 624–629 (1992).
72. D. W. Heinrichs, T. E. Hanlon, W. T. Carpenter, The Quality of Life Scale: An instrument for rating the schizophrenic deficit syndrome. *Schizophr. Bull.* **10**, 388–398 (1984).
73. N. C. Andreasen, *Scale for the Assessment of Positive Symptoms (SAPS)* (University of Iowa, 1984).
74. N. C. Andreasen, The scale for the assessment of negative symptoms (SANS): Conceptual and theoretical foundations. *Br. J. Psychiatry* **155**, 49–52 (1989).
75. W. Guy, "Abnormal involuntary movement scale (AIMS)" in *ECDEU Assessment Manual for Psychopharmacology* (National Institute of Mental Health, 1976).
76. A. Bulat, G. Tzimiropoulos, "How far are we from solving the 2D & 3D Face Alignment problem? (and a dataset of 230,000 3D facial landmarks)" in *2017 IEEE International Conference on Computer Vision (ICCV)* (2017), pp. 1021–1030, 10.1109/ICCV.2017.116.
77. A. Gudi, Recognizing semantic features in faces using deep learning. arXiv [Preprint] (2016). https://doi.org/10.48550/arXiv.1512.00743 (Accessed 9 May 2022).
78. P. Ekman, Universal facial expressions of emotion. *Calif. Ment. Health Res. Dig.* **8**, 151–158 (1970).
79. P. Ekman, Basic emotions. *Handb. Cogn. Emot.* **98**, 16 (1999).
80. R. W. Stineman, A consistently well-behaved method of interpolation. *Creat. Comput.* **6**, 54–57 (1980).
81. S. Moritz, T. Bartz-Beielstein, imputeTS: Time Series Missing Value Imputation in R. *R J.* **9**, 207–218 (2017).
82. S. T. Lane, K. M. Gates, Automated selection of robust individual-level structural equation models for time series data. *Struct. Equ. Model. Multidiscip. J.* **24**, 768–782 (2017).
83. R. V. Lenth *et al.*, emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.8.5. https://CRAN.R-project.org/package=emmeans. Accessed 10 August 2022.
84. T. Hothorn, F. Bretz, P. Westfall, Simultaneous inference in general parametric models. *Biom. J.* **50**, 346–363 (2008).
85. N.T. Hall *et al.*, Data repository: Automating the Analysis of Facial Emotion Dynamics: A Computational Framework and Application in Psychotic Disorders. Open Science Framework. https://osf.io/8gsye/. Deposited 13 March 2024.
86. K. M. Gates, P. C. M. Molenaar, F. G. Hillary, S. Slobounov, Extended unified SEM approach for modeling event-related fMRI data. *NeuroImage* **54**, 1151–1158 (2011).