Edna Bennett Pierce Prevention Research Center
College of Health and Human Development
The Pennsylvania State University
314 Biobehavioral Health Building
296 Henderson Drive
University Park, PA 16802-6505

814-867-2119
prevention@psu.edu

August 3, 2025

Jennifer E. Merrill, PhD
Associate Editor, *Psychology of Addictive Behaviors*

Dear Dr. Merrill:

Thank you very much for the opportunity to revise our manuscript entitled ***Common and Unique Latent Transition Analysis (CULTA) as a Way to Examine the Trait-State Dynamics of Alcohol Intoxication***. We thank the editor and reviewers for their constructive comments on our manuscript.

In this revision, we made several substantive changes to improve clarity, transparency, and analytic rigor. We clarified the distinction between person- and day-level analysis, defined key terms early in the manuscript, and expanded the theoretical justification for our modeling framework. To avoid ambiguity in terminology and align with reviewer feedback, we have revised the labels for the two latent profiles. "High trait HED" is now referred to as "chronic HED" to emphasize its persistent high-intoxication pattern. "State HED" is now labeled "inertia driven drinking" to highlight its episodic and reactive intoxication pattern, characterized by within-person autoregressive dynamics. We added explicit *a priori* hypotheses, equations for the CULTA model, and a Transparency and Openness subsection. We also revised overstated claims, added information on sample demographics and sensor validity, and tempered interpretations based on the study's six-day window.

Furthermore, we completed a Monte Carlo simulation study to evaluate the CULTA model's parameter recovery, convergence behavior, and sample size requirements. The simulation findings are now included in an Appendix and support the CULTA model's ability to reliably recover key parameters (e.g., autoregressive coefficients, profile-specific means, transition probabilities) under realistic conditions.

Together, these changes better align the manuscript with reviewer feedback and enhance its transparency, methodological justification, and relevance to *Psychology of Addictive Behaviors* readers. Point-by-point summary of changes we made are summarized below. Text added or modified in the manuscript in response to the reviews appears in blue.

Comments by Associate Editor (AE)

AE#1: As per the online guidelines, please indicate in your cover letter "a list of published, in press, and under review studies that come from the same dataset as the one in the submitted manuscript, as well as a narrative description of how the submitted manuscript differs from the others." Also note that "upon acceptance of a manuscript, authors will be required to provide, as part of the author note, a list of related published papers that come from the same dataset, unless such papers are clearly described and referenced in the manuscript (specifically noting that findings come from the same dataset)." Despite the different approach to their analysis, if these are the same data used in Russell et al (2024), that should be stated clearly for readers.

We now clarify in both the Author Note and Introduction that the data were used in prior

studies (Russell et al., 2025; Richards et al., 2025), which addressed different questions using distinct analytic approaches. We added the following on p. 2 and 8:

> **"The data reported in this manuscript were previously used in Russell et al. (2025) and Richards et al. (2025). Russell et al. (2025) used multilevel latent profile analysis (MLPA) to identify day-level intoxication profiles based on transdermal alcohol concentration (TAC) features (peak, rise rate, fall rate, and duration) and tested their associations with drinking behaviors, contexts, and the Alcohol Use Disorders Identification Test (AUDIT) scores. Richards et al. (2025) extended the MLPA framework to predict the co-occurrence of negative and positive alcohol-related consequences from these profiles. The present manuscript introduces a novel statistical framework—Common and Unique Latent Transition Analysis (CULTA)—to model both between- and within-person trait- and state-level TAC dynamics and their transitions across days. Unlike prior studies, we model autoregressive inertia, separate common vs. unique variance components, and examine transitions across latent intoxication profiles over time. Thus, the present manuscript addresses different research questions and employs a distinct analytic approach."**

> **"This study uses the same dataset as Russell et al. (2025) and Richards et al. (2025), which previously applied multilevel latent profile analysis (MLPA) to characterize daily intoxication profiles and their associations with drinking behaviors and consequences. In contrast, the present study applies the CULTA model to TAC data collected over six days from young adults engaging in HED. Our primary aim was to characterize both stable and situational components of intoxication and examine short-term dynamic transitions between latent intoxication profiles."**

AE#2: As per the online guidelines, please include a subsection on Transparency and Openness.

We added a Transparency and Openness subsection (Method, p. 18) following APA guidelines. It states:

> **"We report all data exclusions, manipulations, and measures in the study, in accordance with JARS (Kazak, 2018). The study design and analyses were not preregistered. Due to IRB restrictions and participant confidentiality, the raw TAC data are not publicly available. However, all model code and example data are available on OSF (`https://osf.io/gtdmr`) and GitHub (`https://github.com/jeksterslab/manCULTA`, `https://jeksterslab.github.io/manCULTA/index.html`) to enable replication and adaptation of our analytic approach. Analyses were conducted in `Mplus` version 8.11 (Muthen & Muthen, 2017) and `R` version 4.5.1 (R Core Team, 2025). This study complies with the "disclosure" standards of the Transparency and Openness Promotion Guidelines."**

AE#3: In my view, *a priori* hypotheses always strengthen a paper. That said, they should not be stated post hoc. Did the authors have any *a priori* hypotheses that can be stated?

Thank you for highlighting the importance of *a priori* hypotheses. While the CULTA model is a novel framework, our expectations were shaped by prior research on trait–state models, autoregressive drinking dynamics, and latent profile modeling with TAC data. Based on this foundation, we developed three hypotheses prior to model estimation. We have now added these explicitly at the end of the first section of the Introduction (p. 8) to clarify our analytic expectations from the outset. These hypotheses are also aligned with the study's three research questions, which are revisited in detail in the "Research Questions" section. The *a priori* hypotheses are stated as follows:

> **"We tested the following hypotheses: 1) Common vs. Unique Variance: We hypothesized that TAC features would exhibit both common (shared) and unique (feature-specific) variance at both the between- and within-person levels. 2) Latent Profiles: We hypothesized the emergence of at least two distinct latent profiles of intoxication: one reflecting persistently high intoxication, and another reflecting moderate, fluctuating intoxication. 3) Intoxication Inertia and Transition Dynamics: We hypothesized that individuals in the moderate, fluctuating intoxication profile would show significant autoregressive carryover in intoxication levels (i.e., inertia), while those in the persistently high intoxication profile would not. Additionally, we expected baseline Alcohol Use Disorders Identification Test (AUDIT) scores to predict both initial profile membership and the likelihood of remaining in or transitioning between profiles."**

AE#4: Please provide more detailed descriptives on the race and ethnicity distribution in the sample.

We have expanded the participant description in the Method section (p. 15) to include more detailed race, ethnicity, and educational background information. The participant description is stated as follows:

> **"The study involved 222 young adults with an average age of 22.3 years. The sample was 64% female, 79% non-Hispanic White, 84% undergraduates, 6.4% graduate students, and 9.5% nonstudents. Participants were recruited from the vicinity of a northeastern U.S. university using flyers and online postings. No evidence of demographic bias was observed between those who completed the study and those who did not ($ps > .10$)."**

AE#5: The authors acknowledge that the six-day period may not fully capture longer-term patterns or transitions, which was a primary concern as I read the paper. Please review carefully for places where the results/conclusions can be qualified by indicating this short time frame. For example, even the research question might be restated as "Does an individual's baseline AUD influence their initial profile membership or

their likelihood of SHORT-TERM transitions between profiles? Or "Does an individual's baseline AUD influence their initial profile membership or their likelihood of transitions between profiles OVER 6 DAYS? Reviewer 1 agreed that tempering the framing of tests of stability over time was necessary.

AE#6: It seems that day of week is something important to consider, given natural variations between drinking behavior on weekdays vs weekends. Did the authors consider including this in the models?

We appreciate important points raised in comments 5 and 6 and fully agree that the limited six-day window constrains the conclusions we can draw about long-term drinking patterns or profile stability. In response, we carefully revised the manuscript to temper any language implying broader temporal generalizability. Specifically, we now explicitly refer to transitions as "short-term" or "over a six-day period" in the relevant research question, Results, and Discussion sections. We have also added language throughout the manuscript to clarify that our inferences pertain only to dynamics observed within this limited timeframe. The limitations of the short observation period are now discussed more explicitly in the Limitations section.

We added the following research question on p. 15:

> **"Does an individual's baseline AUD influence their initial profile membership or their likelihood of short-term transitioning between profiles over a six-day period?"**

We added the following in the Results section on p. 22:

> **"The results indicate that individuals do shift between high and low intoxication profiles across the six-day observation window, with transition probabilities influenced by both their previous day's profile and their AUDIT scores. These transitions, although limited in scope due to the short study period, provide insight into short-term dynamics of intoxication."**

We added the following in the Discussion section on p. 28 and p. 29:

> **"Transitions between profiles over the six-day period may reflect short-term variability in intoxication patterns, but the limited timeframe precludes conclusions about longer-term stability."**

> **"[T]he study's six-day observation window constrains the ability to examine long-term drinking dynamics or profile stability. While the short duration ensured high compliance and allowed for fine-grained monitoring, all conclusions regarding persistence or transitions must be interpreted as short-term patterns. Future work should employ longer assessment periods (e.g., weeks or months) to assess the robustness of the profiles over time and their susceptibility to external influences."**

"Additionally, day-of-week variation is known to influence drinking behavior. Our study partially addressed this by standardizing the six-day window across participants, covering both weekdays and the typical "social weekend." However, day-of-week effects were not explicitly modeled. Future studies should incorporate cyclical structures to better account for temporal confounds in drinking patterns."

AE#7: It is noted that power analysis could not be conducted. Can the authors provide any indication of whether the sample size was sufficient for the type of modeling conducted here?

We agree that evaluating sample size adequacy is critical given the complexity of the CULTA model. Although we did not conduct an a priori power analysis due to the novelty of the model, we now address this limitation through a dedicated Monte Carlo simulation study.

As detailed in the revised manuscript, the simulation evaluated the CULTA model's performance across five sample sizes ($N = 100, 200, 300, 400, 500$), using population-generating parameters that matched the empirical model. The simulation incorporated autoregressive effects, trait and state variance components, profile-specific means, and covariate-dependent transition parameters—fully reflecting the complexity of the applied model.

Results indicated that CULTA achieved acceptable parameter recovery and inference quality for sample sizes of $N = 200$ and above. Relative bias was generally within ± 0.10, and RMSE values were consistently ideal (< 0.10) or acceptable (0.10–0.20) for most parameters at these sample sizes. Confidence interval coverage was close to the nominal 95% level, and power to detect key parameters exceeded 80% by $N = 200$ and improved further with larger samples. As expected, parameters with small population values near zero (e.g., low log-odds or autoregressive effects) exhibited lower power, reflecting the inherent difficulty of detecting subtle effects.

These findings support the feasibility and robustness of the CULTA model for sample sizes of $N = 200$ or greater and lend confidence to the inferences drawn from our empirical sample ($N = 222$).

Comments by Reviewer 1 (R1)

Thank you for the opportunity to review this manuscript, which examines a novel analytic approach to studying alcohol intoxication. The authors report on CULTA, which merges prior models, common and unique trait-state (CUTS) model and latent transition analysis (LTA). In order to test this novel approach, a sample of young adults ($N = 222$) provided six days of transdermal alcohol concentration (TAC) data. Analysis revealed two latent intoxication profiles, high trait HED (persistent high intoxication) and state HED (moderate episode intoxication. They also revealed that higher AUDIT predicted the high trait HED profile. Individual TAC features were also examined on variability. Overall, the manuscript presents a novel and interesting method for understanding drinking patterns with objective fine-grained data. However, there are several important concerns regarding the background, appropriateness of the data to answer the stated questions, as well as the conclusions drawn. Several specific points of feedback are described below.

R1#1: Heavy episodic drinking versus heavy alcohol use should be defined early in the introduction for the reader, as there is some inconsistency in the literature in the use of these terms and frequent misunderstanding.

To improve clarity and consistency, we have moved the definitions of heavy episodic drinking (HED) and heavy alcohol use from a footnote into the main body of the text early in the Introduction (p. 5). These terms are now explicitly defined in accordance with the SAMHSA (2023) criteria and are used consistently throughout the manuscript. The paragraph of the manuscript now reads as follows:

> **"Alcohol consumption among young adults remains a critical public health concern. According to the 2022 National Survey on Drug Use and Health, 50.2% of individuals aged 18 to 25 reported drinking in the past month, with 29.5% engaging in heavy episodic drinking (HED)—defined as consuming five or more drinks for males or four or more drinks for females on a single occasion at least once in the past 30 days—and 7.6% engaging in heavy alcohol use, defined as engaging in HED on five or more days within the past 30 days, following the previously stated thresholds for males and females (SAMHSA, 2023). Alcohol-related incidents contribute significantly to morbidity and mortality in this age group, with an estimated 1,519 deaths annually among college students aged 18 to 24, and an additional 2,586 deaths from alcohol-related injuries, such as motor vehicle accidents (Hingson et al., 2017). In addition to health risks, alcohol misuse carries substantial economic costs, putting a strain on the healthcare and social systems (Sacks et al., 2015)."**

R1#2: It is a bit unclear at times when reading the introduction as to whether the characterization of drinking is at the person-level or the day-level. Given the analysis integrates between- and within-sources of variance, this may be a difficult task, but I suggest trying to make the ultimate level of analysis and implications of them clear early in the framing of the question.

In the revised manuscript, we now state explicitly at the beginning of the Introduction that our characterization of drinking spans both person-level and day-level variability. Specifically, we have clarified that transdermal alcohol concentration (TAC) features are measured at the day level, and that the CULTA model is designed to decompose variance in these daily features into between-person (trait) and within-person (state) components. We also note the distinction between these levels of analysis and how they relate to our research questions and interpretations. We added the following on p. 9:

> **"In this study, we characterize alcohol intoxication at the level of daily TAC features, which reflect drinking events within a 24-hour "social" day. Our analytic approach integrates both day-level (within-person) and person-level (between-person) sources of variability using the CULTA model. This framework allows us to distinguish stable, trait-like tendencies from situational, state-level fluctuations in intoxication across drinking days. Clarifying this distinction is essential for understanding the**

implications of our findings, which address both persistent individual differences and dynamic transitions in drinking behavior over time."

R1#3: It is a bit unclear what it might mean to account for "lingering but dissipating effects of intoxication" with these type of data, given many readers may considering subjective forms of intoxication, which are not studied here. I would suggest clarifying these constructs early, as they are more clearly defined in the methods.

We agree that the phrase "lingering but dissipating effects of intoxication" could be misinterpreted as referring to subjective intoxication, which is not measured in our study. In the revised manuscript, we now clarify that this phrase specifically refers to the autoregressive effect modeled in the CULTA framework, which captures statistical carryover in latent state intoxication from one day to the next. In other words, "lingering" refers to the degree to which intoxication on a given day predicts intoxication on the following day—above and beyond a person's trait level—while "dissipating" reflects the decay of this effect toward the stable mean over time. This is implemented in the model as an AR(1) process on the latent state intoxication factor, rather than any particular TAC feature or subjective report. We have clarified this conceptualization in the Introduction and revised the relevant sentence on p. 9 as follows:

> **"To effectively capture the complex, time-dependent dynamics of alcohol intoxication, the CULTA model extends the CUTS model's focus on common and unique variability (Hamaker et al., 2016) to incorporate LTA as a means of identifying high- and low-risk intoxication days and examining the shifts in these intoxication patterns over time. CULTA achieves this by distinguishing stable, trait-like components from transient, state-like variations, accounting for the lingering but dissipating effects of intoxication. In this context, these effects correspond to the autoregressive component of the CULTA model, which captures the extent to which a person's prior-day state intoxication level predicts their current-day level, with influences that diminish over time toward their average."**

R1#4: The background acknowledges that environmental factors may impact these state-trait patterns over time, yet does not discuss the value of the research question in the sample studied (i.e., young adults). While the background does a thorough job of identifying what about the method is valuable and novel, it is less clear what the implications might be specifically for the sample.

We agree that the background would benefit from a clearer articulation of why this research question is particularly relevant for young adults. In the revised manuscript, we have expanded the Introduction to emphasize that young adulthood is a developmental period characterized by high variability in drinking behavior, increased risk for heavy episodic drinking, and elevated susceptibility to alcohol-related harms. These characteristics make young adults an especially informative group for examining both trait-like persistence and state-like fluctuations in intoxication.

We now explicitly state that understanding how stable versus transient intoxication patterns

manifest in this population may inform early intervention efforts, given that drinking habits during this period often lay the foundation for longer-term trajectories of alcohol use and risk. We have added the following sentences to the Introduction (p. 5):

> **"Young adulthood is marked by heightened contextual reactivity, variability in drinking behavior, and elevated risk for both acute and long-term alcohol-related consequences (Arnett, 2005; Patrick & Terry-McElrath, 2016; Schulenberg et al., 2003). Examining intoxication dynamics in this developmental period may provide insight into early patterns of risk persistence or escalation, and help identify which individuals are more likely to exhibit inertia in heavy drinking versus those who fluctuate in response to situational influences. These distinctions have important implications for the timing and tailoring of interventions (Maggs & Schulenberg, 2005)."**

R1#5: The implications of the findings may be slightly overstated. Specifically, the statement that the results "illuminate the mechanisms underlying persistent heavy drinking", as the analyses seem to more clearly characterize patterns of drinking, rather than study 'mechanisms' leading to those patterns.

We agree that our use of the term 'mechanisms' may overstate what can be inferred from our analytic approach. While the CULTA model allows us to distinguish persistent from transient intoxication patterns and estimate autoregressive carryover in state-level intoxication, these patterns are best described as characterizations of drinking dynamics rather than direct evidence of underlying psychological or neurobiological mechanisms.

To address this concern, we have revised the sentence on p. 22 to better reflect the nature of our contributions:

> **"These findings contribute to a nuanced characterization of intoxication dynamics and may offer insight into patterns of persistence and variability in heavy drinking behavior."**

This revision avoids causal language while preserving the value of the analytic distinctions made possible by the CULTA model.

R1#6: Similarly, some broad statements at the end of the introduction could be more directly stated with relevance to the present work, for example "significant implications for public health and behavioral science" and reference to "vulnerable populations" both seem broad and ambitious relative to the potential findings.

We agree that these broad statements at the end of the Introduction could be made more specific to the scope and contributions of the current study. In response, we have revised the closing sentences of the Introduction (p. 5) to more directly tie the potential implications to our study's focus on young adults with a history of heavy episodic drinking, and to ground the

discussion in the observed short-term dynamics. We have replaced the original sentence with the following on p. 22:

> **"By capturing short-term dynamics in intoxication among young adults engaged in heavy episodic drinking, this study provides a foundation for more personalized models of risk and may inform future research on early intervention targets in this developmentally sensitive population."**

We hope these revisions align the tone of our conclusions more closely with the scope of our findings.

R1#7: One of the more significant limitations of this work is that the relatively short longitudinal timespan of the study limits the ability to really test the novel questions posed, namely the stability of these drinking patterns over time. This is likely highlighted in the two-profile solution with stable patterns. While this is included as a limitation of the study, authors may consider tempering the framing further.

We fully agree that the short six-day observation window limits our ability to draw conclusions about long-term stability in drinking patterns or transitions between profiles. In response to the Associate Editor's earlier comment, we have already revised the framing of our research questions and results to explicitly characterize transitions as "short-term" and to clarify that our findings reflect patterns observed only over a six-day period.

To further temper our conclusions, we have made additional edits in the Discussion to emphasize that the persistence of patterns within the two-profile solution may reflect short-term regularities rather than long-term stability. We also now note that the observed stability in the chronic HED profile may reflect a rapid return to profile-specific means within this limited time span rather than enduring trait-like behavior across longer intervals. Specifically, we added the following clarifying sentence on p. 21:

> **"Given the six-day study window, the apparent stability of the chronic HED profile should be interpreted as short-term regularity in intoxication dynamics, rather than definitive evidence of enduring drinking traits. Longer assessment periods are needed to determine whether these patterns persist across weeks or months."**

We hope these revisions make clear that our interpretations are appropriately bounded by the study's timespan.

R1#8: The implications discussed in results may also be slightly over-interpreted. For instance, given the finding that fall rate and drinking duration were the most persistent features, authors conclude that these might be the most impactful intervention targets. However, it is not clear how 'stability' of a drinking features over this short timeframe is indicative of a malleable target? Further, it is not clear how fall rate may be intervened on, given it is a direct result of the other factors of the drinking episode (e.g., rise rate and peak TAC).

We agree that our interpretation of the intervention implications for fall rate and duration could be clarified and tempered. Our intention was not to suggest that the stability of these features over six days alone provides evidence of their modifiability, but rather to highlight that these features exhibited meaningful between-person variability, which may reflect more enduring behavioral tendencies worth further investigation.

To avoid over-interpretation, we have revised the Discussion to make clear that the identification of fall rate and duration as potentially relevant targets is hypothesis-generating rather than prescriptive. We now explicitly acknowledge that fall rate is a physiological consequence of earlier drinking behavior (e.g., rapid or high-volume consumption) and may be better understood as an outcome of modifiable upstream behaviors (e.g., drinking pace, beverage strength, and timing of cessation). We have revised the relevant sentence on p. 20 as follows:

> **"Although fall rate and duration exhibited the greatest between-person variability across the study window, we do not interpret this as evidence of modifiability *per se*. Rather, their persistence suggests that they may reflect broader individual differences in drinking style or episode structure. In particular, fall rate likely reflects downstream effects of earlier drinking behaviors (e.g., rapid consumption or delayed cessation), and may therefore be indirectly influenced through interventions targeting these upstream factors."**

We hope this revision better communicates the nuance of our interpretation and avoids overstating the intervention implications of short-term stability in these features.

R1#9: The label of the low profile as "state HED" is slightly misleading as these are events that are characteristically low drinking. While I understand the inclusion criteria for the study leads the authors to assume this group still includes individual (i.e., state HED) episodes, it is not clear that the data is reflecting this type of pattern and attempted replication of these methods may be misled by this naming.

We agree with the reviewer's concern that the prior label "state HED" could have been misinterpreted as implying uniformly high intoxication levels across all days within the profile. To address this, we have renamed the profiles to clarify their conceptual distinctions while avoiding potential misinterpretations. The profiles are now labeled as:

- **Chronic HED: This profile reflects individuals who exhibit persistently elevated intoxication across days, indicative of a more stable, high-risk drinking pattern.**
- **Inertia driven HED: This profile reflects individuals who exhibit moderate, episodic intoxication with significant autoregressive carryover, meaning intoxication levels on a given day are statistically influenced by the prior day's levels, consistent with inertia dynamics.**

This renaming provides a clearer, behaviorally grounded interpretation of the profiles, emphasizing that "inertia driven drinking" captures state-like, fluctuating intoxication with

lingering effects, while "chronic HED" reflects sustained high intoxication levels without implying direct equivalence to subjective states or specific drinking events. We added the following on p. 20 to clarify this point:

> **"The high profile, which we labeled as chronic HED ($c = 0$), was associated with systematically elevated mean values across the TAC features, indicating more intense and prolonged intoxication episodes relative to the sample average. Although individuals can transition between profiles, when in the chronic HED profile, individuals exhibit a persistent configuration of elevated intoxication features that differs meaningfully from typical patterns observed in the sample. This profile is *trait-like* in the sense that the mean structure of intoxication features remains consistently elevated within each day, independent of prior-day state. This reflects a stable intoxication expression characteristic of this profile, as evidenced by the non-significant AR parameter (ϕ_0), indicating little carryover from previous-day intoxication."**

> **"In contrast, the low profile, which we labeled as inertia driven drinking ($c = 1$), corresponds to lower mean values for the same TAC features, suggesting milder or shorter intoxication episodes relative to the high profile. Individuals in this profile exhibit lower peak intoxication levels and more moderate drinking patterns, with shorter episodes and quicker changes in alcohol concentration. This profile represents the "typical" drinking behavior within the sample, which, as noted earlier, is composed of individuals who engage in HED. Therefore, typical in this context refers to less intense but still episodic drinking behavior, which may be common among the majority of participants. We labeled this profile inertia driven drinking to emphasize its episodic and reactive nature, in contrast to the persistent elevation observed in the chronic HED profile. This profile is *state-like* because intoxication levels fluctuate depending on the previous day's state, consistent with significant autoregressive carryover ($\phi_1 = 0.311, p < 0.001$). These fluctuations eventually dissipate, with intoxication levels tending to return to the profile's lower mean configuration over time."**

To summarize the core distinctions between trait- and state-like dynamics within the CULTA framework—and to contrast with Research Question 1's focus on the decomposition of variance—Table 5 provides an overview of each model component, its operationalization, and its substantive interpretation.

Comments by Reviewer 3 (R3)

I'd like to congratulate the authors on a really cool project and study. The combination of latent transition modeling and the CUTS model is very promising and can give novel insights into many processes. Despite this I do have some concerns with the paper that need to be addressed or discussed in more detail. My four main concerns are:

R3#1: There is no information about the reliability and the validity of the TAC sensor or of the measures derived from them. Without this information it is impossible to judge the strength of results accurately. This reliability and validity id especially important because the TAC features are not directly derived from dedicated sensors but are derived quantities based on sensor input of something else. Also the type of sensor used will be validated on young white people, which if a further threat of validity for measures obtained from people with older or darker skin among other things.

We revised the Method section to include more information about the reliability and validity of the TAC sensors and derived features. Specifically, we now note that the SCRAM CAM sensor used in the present study has been validated in multiple controlled laboratory studies (e.g., Roache et al., 2019) and is widely used in both clinical and naturalistic settings. The core TAC signal is reliably recorded, and derived features (e.g., peak, rise rate, fall rate, duration) have been shown to correlate strongly with self-reported drinking quantity and predict alcohol-related outcomes even after controlling for drink counts (Russell et al., 2022; Fridberg et al., 2022).

We also acknowledge the important limitation regarding generalizability of sensor calibration. Most validation studies for the SCRAM CAM device have been conducted on predominantly non-Hispanic White participants. As such, we have now added to the Discussion section a statement on the potential limitations of skin-based alcohol detection due to variation in skin tone, temperature, hydration, and age, all of which may affect the transdermal signal. Although we found no evidence of differential compliance or missingness by race/ethnicity in this sample, we agree that further work is needed to ensure equitable measurement validity across populations. We now explicitly flag this as a limitation and a direction for future work on p. 28:

> **"[A]lthough TAC sensors provide objective and continuous alcohol monitoring, their validity may vary across populations. Most validation studies for the SCRAM CAM device have been conducted in samples that are predominantly non-Hispanic White (Fridberg et al., 2022; Roache et al., 2019; Russell et al., 2022). This raises concerns about transdermal signal accuracy across individuals with different skin tones, hydration levels, body temperatures, or ages. Although we found no evidence of differential compliance or missingness by race/ethnicity in this study, future research is needed to evaluate measurement validity in more diverse populations."**

R3#2: The way that the CULTA model is setup it doesn't allow the detection of different states in the common intoxication construct or the unique aspects of the TAC features. It merely allows for between class differences in the latent intercepts of the observations and the AR effect. This does not imply between class differences in the common and unique constructs however. It merely makes class membership an additional source of variation in observed scores next to differences in the common construct etc. Also, the lack of formulas for the CULTA model (especially for the state and state transition part make the model hard to understand. The reference to the figures that are made on page 8 are not enough.

We appreciate this thoughtful and detailed comment, and we are grateful for the opportunity to clarify both the theoretical structure and the implementation of the CULTA model. In the

revised manuscript, we have added a sequence of formal model equations and explanatory text to explicitly define the measurement, dynamic, and transition components of the CULTA framework. Specifically, we now include:

- **the original CUTS measurement model (Equation 1),**
- **the CULTA measurement model with profile-specific means (Equation 2),**
- **the autoregressive dynamics governing the latent state intoxication process (Equation 3), and**
- **the logit-based initial and transition models for latent profile membership (Equations 4 and 5).**

These additions clarify how CULTA integrates the CUTS model with latent transition analysis and make the mechanisms of the model transparent for the reader.

We agree with the reviewer that profile membership is introduced as an additional source of systematic variation beyond the common trait and common state processes. The profile-specific means ($\mu_{k,c}$) are estimated at the level of the observed indicators, and it is true that they do not directly moderate the latent trait or latent state variables. However, these parameters are not arbitrary; they reflect systematic structure in the unique variances—those not captured by the shared trait and state components. In other words, after the shared components of each TAC feature are removed (both between and within-person), the profile structure captures stable configurations in the remaining, feature-specific uniquenesses.

Thus, while centering is indeed a part of what these profile-specific parameters do, it is not merely re-centering in a statistical sense. Rather, the profile means reflect emergent profiles of intoxication that persist even after accounting for shared dynamics. When these means are similar across indicators, they reflect an overall shift in latent intoxication levels; when they differ, they suggest that certain features contribute uniquely to the configuration of that profile.

From this perspective, transitions between latent profiles can be interpreted as changes in the unique profile of intoxication expression—shifts in how the pattern of feature-specific deviations evolves over time. This is substantively meaningful because it points to behavioral configurations that are not reducible to trait- or state-level intoxication alone. We have revised the manuscript to better explain this distinction and its implications for interpretation.

We added the section Interpretation of Profile-Specific Means to summarize these points on p. 13.

> **"In the CULTA measurement model (Equation 2), profile-specific means ($\mu_{k,c}$) are introduced as intercept shifts at the level of the observed TAC features. While these parameters do not directly moderate the latent trait or latent state variables, they are**

not arbitrary constants. Rather, they reflect systematic structure in the residual variances—i.e., the parts of the TAC features not explained by the shared trait and state components. Once the common components of each TAC feature are removed (both between- and within-person), the profile structure captures persistent patterns in the remaining feature-specific uniquenesses. These patterns are not merely statistical artifacts but convey meaningful and stable differences in intoxication expression. From this perspective, $\mu_{k,c}$ parameters can be interpreted as emergent profiles that describe how unique aspects of intoxication differ across latent profiles. When profile-specific means are similar across TAC features, they reflect an overall shift in the expression of intoxication; when they diverge, they reveal differential contributions of features such as peak, rise, fall, or duration to each latent profile. Consequently, transitions between latent profiles do not just indicate movement across broad intoxication levels—they also represent changes in the configuration of intoxication expression. Such transitions are substantively meaningful because they reflect behavioral profiles that are not reducible to trait- or state-level intoxication alone."

We also added Table 3 summarizing all model parameters and their substantive interpretations. It provides a compact reference that distills the content from the preceding sections on the model structure, intoxication inertia, and profile transitions.

Finally, we enhanced the narrative and visual explanation by revising the figures and adding extensive interpretive text surrounding Figures 2 and 3 to guide the reader through the model logic. These revisions help clarify that CULTA is not merely a mixture model with profile-varying intercepts, but a framework capable of modeling substantively meaningful transitions between latent state configurations, with important implications for identifying short-term drinking risk dynamics.

We hope these revisions resolve the concerns regarding transparency, interpretability, and model capability.

R3#3: No information is given in the article about the fit of the CULTA model used for results. The fact that many of the sources of variance in the model are not significant seem to imply that the model does not fit the data well. The fact that there are only 2 recovered classes, one "high" and on "low" class, further enforces this suspicion since a LCA will tend to always give this distinction when there is not really anything to discover. The model wants to separate classes somehow, and "in lack of anything better" it will go for a distinction based on high and low scores. In addition, the results here could also be explained, for example, simply by the fact that i) after an episode of heavy drinking, people they tend to drink less at the next measurement (negative AR in drinking), and ii) not everyone has episodes of heavy drinking. Then your two classes merely reflect that not everybody engages in heavy drinking (this is where the high vs low distinction comes from, a simple mean difference), and the transition probabilities just reflect the negative AR of heavy drinking. Both results already known from "simple" AR models. The lack of an AR effect in the high group finally could simply be a ceiling effect. Point is, that there are many possible explanations/models for the patterns of results you present, and without proof that the model fits the data well, results from this model cannot be judged properly. If the

model does not fit well, interpreting it doesn't make sense after all. This also means that all the conclusions and recommendations drawn based on the data analysis could be premature (also because of the next point).

We agree that establishing model fit is essential for evaluating the validity of our conclusions. In the revised manuscript, we now provide detailed information about model fit indices for the CULTA model, including the log-likelihood, number of parameters, AIC, BIC, and sample-size adjusted BIC. We also report entropy to support the adequacy of classification quality in the latent transition component. These values are presented in Table 3 (p. 37).

We interpret the lack of significant variance in some components (e.g., ψ_{peak}, ψ_{rise}) not as evidence of model misspecification, but as indication of relative homogeneity in those features across individuals. The significant common and unique state variances, along with unique trait variance in fall rate and duration, support the conceptual distinction between trait-like and state-like aspects of intoxication within the CULTA framework.

We recognize that high vs. low distinctions are common in latent profile analysis. However, our profile enumeration process considered models with one, two, and three profiles. The two-profile model was selected based on superior information criteria, classification quality, and interpretability. Three-profile models encountered convergence issues and failed to yield distinct or stable profiles. Thus, we did not impose a high-low dichotomy but selected the most parsimonious and empirically supported model.

> **"[C]ertain technical limitations of the CULTA model warrant further discussion. We encountered convergence issues when estimating models with more than two latent profiles, suggesting either limited data support or increased model complexity. While the two-profile solution (chronic HED vs. inertia driven drinking) was theoretically interpretable and fit the data well (as supported by AIC, BIC, and entropy), it may not capture the full spectrum of heterogeneity in intoxication patterns. Larger and more varied samples may allow for more complex profile structures."**

> **"A potential limitation of this study concerns the interpretation of model fit and latent profile structure. Although we reported detailed model fit indices—including log-likelihood, AIC, BIC, adjusted BIC, and entropy—that indicated good model performance, several trait-level variance components were not statistically significant. We interpret this as evidence of relative homogeneity rather than model misspecification, but we recognize that null variance estimates can raise concerns about overparameterization or limited sensitivity. Importantly, trait-level variance in fall rate and duration remained significant, suggesting meaningful between-person differences in those features."**

The near-zero AR effect in the chronic HED profile reflects a statistical consequence of ceiling effects, not an absence of behavioral persistence. Because individuals in this group maintain consistently high intoxication levels near the upper measurement limit, there is little room for

day-to-day variability. AR models capture predictability in fluctuations around the mean, so when variability is constrained, deviations appear random, reducing the AR coefficient. Thus, the lack of AR in chronic HED indicates rigid, stable heavy drinking with minimal short-term fluctuations, not low persistence. We added the following on p. [XXX] in the Limitations section to clarify our interpretation of thr results.

> **"An important consideration in interpreting the present findings involves the potential influence of ceiling effects on the observed autoregressive (AR) dynamics within the chronic HED profile. Ceiling effects occur when measured values approach the upper bounds of a scale, limiting observable variability and potentially obscuring temporal dependencies such as autoregressive carryover. In our study, individuals classified in the chronic HED profile exhibited consistently elevated mean levels across all TAC features, including peak intoxication, rise rate, fall rate, and duration. This high, persistent intoxication pattern inherently restricts the range for day-to-day fluctuations, compressing variability near the upper end of the measurement scale."**

> **"From a statistical perspective, AR models capture systematic predictability in fluctuations around the mean, not the mean level itself. Thus, when variability is minimal and constrained by ceiling proximity, any residual deviations from the mean tend to appear as random noise. This results in an AR coefficient that approaches zero—not because inertia is absent, but because the limited fluctuations lack structured, time-dependent patterns. Importantly, this statistical phenomenon aligns with the substantive interpretation of entrenched, inflexible drinking behavior in the chronic HED profile."**

> **"Rather than viewing the absence of a significant AR effect in the chronic HED profile as an indication of low persistence, we contend that it reflects a different, equally concerning behavioral dynamic. The clustering of data points near the ceiling, combined with minimal random variability, suggests that individuals in this profile have stabilized at consistently high intoxication levels. Their drinking behavior no longer exhibits meaningful reactivity to day-to-day influences but instead demonstrates rigidity, characterized by persistent heavy intoxication largely unaffected by situational fluctuations."**

> **"This pattern has clear clinical implications. The chronic HED profile appears to represent individuals with high-risk, habitual drinking patterns that resist contextual modulation. In contrast, individuals in the inertia-driven drinking profile show moderate intoxication levels coupled with significant AR dynamics, reflecting episodic intoxication with lingering effects that gradually dissipate over time. Together, these profiles highlight distinct pathways to alcohol-related risk: one rooted in entrenched, persistently elevated intoxication with minimal fluctuation, and the other in fluctuating, reactive drinking patterns with carryover effects."**

"Moving forward, these findings underscore the importance of considering ceiling effects not merely as statistical artifacts but as meaningful indicators of behavioral inflexibility. Future work should replicate these patterns over longer observation windows and explore alternative modeling approaches—such as censored or nonlinear time series models—that explicitly account for bounded measurement scales. Such efforts will further clarify the interplay between intoxication dynamics, ceiling effects, and risk for sustained heavy drinking."

R3#4: Since you're introducing a new method a simulation study is important. To show that it can adequately recover the underlying process. Now we have no idea how well the CULTA model performs and the application can't tell us that either since we don't know what the true process looks like.

We agree that a simulation study is essential to evaluate the CULTA model's reliability and performance. In response, we added a dedicated Monte Carlo simulation section to the manuscript. This simulation used data generated from the full CULTA model—including trait-state decomposition, profile-specific inertia, and covariate-influenced transitions—and compared CULTA against LTA and RILTA, which omit key structural components.

Across five sample sizes ($N = 100$–500), CULTA demonstrated superior model fit, improved parameter recovery, and more valid inference than the comparison models. Parameters were recovered with minimal bias and acceptable RMSE by $N = 200$, and statistical power exceeded 0.80 for most nonzero parameters at that sample size. As expected, parameters near zero showed lower power. Confidence interval coverage was close to nominal, and convergence rates were high across all conditions. These results support CULTA's robustness under realistic conditions and confirm the adequacy of our empirical sample ($N = 222$) for reliable estimation and inference.

We agree that further validation work is essential for advancing CULTA as a general-purpose modeling tool.

I also have a few smaller point which I will list below:

Page 6: "As TAC measures are highly correlated... most important and meaningful variance".

Even in multiple regression this overlap would still be taken into account and would show up in the R-squared of the model, even if not in the regression coefficients.

We agree that in multiple regression, shared variance is still reflected in model R^2, even if not in individual coefficients. Our point here was not to imply that overlap is lost, but rather to motivate the use of a latent variable approach that can formally partition shared (common) and unique variance—something that regression does not do explicitly. We have revised the section on p. 6 to avoid confusion and now clarify the advantage of modeling shared variance directly via latent factors.

"Another approach has been to use multiple "elementary" TAC features (e.g., peak,

rise rate, fall rate, duration) in a multiple regression, in which their individual/unique contributions (e.g., Richards, Glenn, et al., 2024) and/or their statistical interactions are tested (e.g., Simons et al., 2015). Testing the unique predictive value of each TAC feature is important, and we acknowledge that in multiple regression models, the overall shared variance among predictors is still retained in the total variance explained (i.e., R^2). However, such models do not explicitly represent or isolate this shared variance, and thus cannot distinguish between the predictive power of common versus feature-specific components. As TAC features are often highly correlated (see, e.g., Russell et al., 2022), modeling them independently may obscure the contribution of the shared intoxication signal they reflect. In contrast, latent variable models can formally decompose TAC features into common and unique sources of variance, allowing clearer interpretation of their joint and separate contributions to alcohol-related outcomes. Moreover, although interaction terms can be informative (Simons et al., 2015), they become increasingly cumbersome, underpowered, and difficult to interpret as the number of predictors grows. The latent trait–state modeling approach used here addresses these challenges by offering a principled way to disentangle overlapping and feature-specific components while capturing dynamic patterns over time."

Page 7: "Models that can parse the common and unique..."

Here you seem to apply a factor model for "drinking". It would be good to make that explicit immediately and to discuss why that makes theoretical sense. It it really the case that there is an underlying "drinking" factor or is it just about how often you engage in certain drinking behaviors?

We appreciate this observation and have now explicitly stated that we are using a dynamic factor analytic framework to model intoxication as a latent construct. Theoretically, this is justified by the interpretation of TAC features (e.g., peak, rise rate, fall rate, duration) as observable indicators of an underlying physiological process—alcohol intoxication. This aligns with prior research using TAC as an objective index of alcohol consumption intensity. We now include a rationale for this measurement model in the revised text on p. 7.

"We propose a dynamic factor analytic framework in which TAC features serve as observable indicators of a latent intoxication construct. This approach is grounded in the idea that features such as peak, rise rate, fall rate, and duration jointly reflect the underlying physiological process of alcohol intoxication, rather than capturing entirely distinct aspects of behavior. Prior research has shown that these TAC-derived features are interrelated and correspond closely with both subjective and behavioral indices of alcohol exposure (e.g., Russell et al., 2022). By modeling intoxication as a latent factor, we can formally partition shared variance (reflecting the common intoxication signal) from unique variance (feature-specific information) and examine how these components vary within and between individuals over time. Models that can parse the common and unique contributions of TAC features simultaneously could be highly

informative in prevention efforts by helping us answer questions such as "should we focus solely on reducing overall intoxication?" or "should we focus on reducing only specific attributes (e.g., speed, duration, intensity) of a drinking episode?""

Page 7: "A framework that can incorporate..."

A multilevel model would also be able to do this, so that's not a gap that's uniquely tackled by the CULTA model.

We agree that multilevel models can also decompose within- and between-person variability. We have revised p. 8 to clarify that CULTA can be viewed as a special case of a multilevel structural equation model. The added value of the CULTA specification lies in (a) modeling the common and unique structure via latent variables, (b) incorporating autoregressive dynamics at the state level, and (c) allowing both AR and intercept parameters to vary across latent profiles, which traditional multilevel models do not typically accommodate in a single framework.

"While standard multilevel models can decompose within- and between-person variability, CULTA can be viewed as a special case of a multilevel structural equation model that extends this framework in key ways. Specifically, CULTA models the common and unique variance across TAC features using latent variables, incorporates autoregressive dynamics at the state level, and allows both intercept and AR parameters to vary across latent profiles. These extensions enable CULTA to simultaneously capture heterogeneity in level and inertia of intoxication across unobserved subgroups, which traditional multilevel models do not typically accommodate in a unified structure."

Page 8: "In the current application of the CUTS framework..."

How do you deal with the fact that the 4 features probably are not normally distributed? And what is the impact of that?

We acknowledge that TAC features are positively skewed. Although the model assumes normally distributed residuals at the indicator level, the use of robust maximum likelihood (MLR) in Mplus helps mitigate bias due to moderate non-normality. We also smoothed the raw TAC data before calculating the four features, which further stabilized their distributions. The potential impact of skew and the robustness of our estimation approach is acknowledged as a limitation on p. 30.

"Another limitation concerns the distributional properties of the TAC features. Several of the derived indicators (e.g., peak, rise rate, duration) were positively skewed, which may violate the assumption of multivariate normality underlying the CULTA model's latent variable structure. Although we applied data smoothing procedures prior to feature extraction to reduce measurement noise and stabilize

distributions, the features remained moderately non-normal. To address this, we used robust maximum likelihood estimation (MLR) in Mplus, which adjusts standard errors and fit statistics to account for non-normality in the observed variables. Nevertheless, violations of normality may still influence model fit or the precision of parameter estimates, and future applications may benefit from exploring alternative distributional assumptions or transformations for highly skewed indicators."

Page 9: The decomposition in the unique traits and the class specific means should also be presented in equations. Right now, all necessary equations for the LTA are not presented.

Please refer to the response to comment #2.

Page 13: The period of data collection seems to include the emergence of COVID-19, how will that have affected data quality?

We completed all data collection prior to the onset of the COVID-19 pandemic. Although our original plan was to collect a larger sample, data collection was halted when the university shut down. As a result, COVID-19 should not have affected data quality.

Page 14: "Prior to analysis, TAC data were smoothed"

How? What were the precise specifications of the b-splines? And why did you use these smoothing approach and how could it have effected results and the validity of the data?

We added to following on p. 17 to describe the smoothing approach:

"To prepare the data for analysis, we applied penalized b-spline smoothing in `PROC TRANSREG` (Eilers & Marx, 1996; Russell et al., 2022). This method effectively smooths noise in complex TAC trajectories while minimizing oversmoothing through derivative-based penalization of spline coefficients. Each TAC drinking episode was modeled separately. We opted for penalized b-splines over simpler filters (e.g., moving averages), which can oversmooth data and obscure meaningful variation (Chatfield, 2003). Although oversmoothing may be less of a concern with high-density sensors (e.g., those collecting data every 20 seconds), SCRAM-CAM captures TAC at 30-minute intervals. Given our goal of extracting features such as peak, rise rate, and fall rate, we sought a method that retained informative signal while mitigating the influence of white noise. Penalized b-splines provided a suitable balance for our sensor resolution and analytic goals."

Page 14: Do you have psychometric information about the AUDIT scale? Has it been validated? For what populations? What is the reliability in this study?

We now include psychometric information on the AUDIT. It has been validated extensively across clinical and non-clinical samples, including young adults (e.g., Reinert & Allen, 2007). In

our sample, internal consistency was acceptable (Cronbach's $\alpha = .80$).

Page 14: "For RQ 1, ..."

How did you test the significance of the variance components? Because the typical t-tests and z-test given by statistical packages are not appropriate. You need to test them with a chi-square test.

We tested the significance of variance parameters by comparing nested models with variance component freely estimated and with the variance component fixed to zero using likelihood ratio tests (LRTs) with chi-square difference testing and scaling correction for MLR.

Page 15: Why did you fix the loading of the "peak" feature to 1? Is this the best, most reliable indicator of the underlying construct? Why not one of the other traits, especially since peak did not have a significant amount of unique trait variance.

We fixed the loading of "peak" to 1 for model identification, not because it is necessarily the strongest indicator. This is standard practice in factor analysis. We now clarify this decision and acknowledge in the Discussion that the unique trait variance for peak was not significant, which is an interesting result—but this does not invalidate its use as the scaling indicator. Using another feature (e.g., fall) was used for identification in another run but this did not significantly change the results.

Page 17: You use scores on the TAC features to identify the latent classes, but 2 of these show no unique trait variance? How does this relate to identify group-based differences between people based on these indicators.

We appreciate this observation. Although two features had non-significant unique trait variance, this does not invalidate their role in latent profile formation. The profile-specific intercepts ($\mu_{k,c}$) capture differences across profiles even when between-person variance is low. We clarify in the Discussion on p. 26 that:

> **"Although peak and rise rate did not show significant unique trait variance, this does not imply they were unimportant for latent profile formation. Because profile-specific intercepts ($\mu_{k,c}$) allow mean-level differences across latent profiles, TAC features can contribute to differentiating profile membership even when individual variability in their trait components is minimal. This highlights how latent profiles can reflect distinct intoxication profiles based on all four TAC features, not just those with high between-person variability."**

Page 19: Here you mention state transitions, how long did people stay in the state they transitioned to? Did they really transition to a high "drinking" state, or did they just have one occasion of drinking a lot. That's an important difference.

We agree that distinguishing between a one-off high-drinking event and a meaningful state transition is essential for interpreting the model's results. As shown in Figure 7, the majority of

individuals in the sample remained in the second profile (inertia driven drinking) across the six-day observation window, suggesting relatively stable moderate intoxication patterns. However, there were individuals who exhibited sustained membership in the first profile (chronic HED) across multiple consecutive days—for example, profile sequences such as 111112 and 111122, while rare, occurred in 0.9% and 0.5% of the sample, respectively. These cases indicate that for a small subset of participants, high intoxication was not merely a single event but a persistent pattern within the timeframe of the study.

We acknowledge that most transitions into profile 1 were short-lived, likely representing isolated heavy drinking episodes. However, the presence of individuals who remained predominantly in profile 1 suggests that CULTA captured not only transient state-level fluctuations but also short-term stability in high-risk profiles, consistent with the notion of trait-like persistence. While the brief six-day period limits our ability to draw conclusions about long-term behavioral patterns, we believe the model was successful in identifying both enduring and episodic intoxication dynamics within the constraints of the data. We made revisions on p. 22 to make these ideas more explicit.

Page 20: When discussing the effect of AUDIT scores on transitions you mention a difference between high and low AUDIT groups, was this verified with a significant interaction/a main effect of AUDIT on transitions?

Yes, the effect of AUDIT on transition probabilities was explicitly tested by including it as a predictor in the multinomial logistic regression for latent profile transitions. The resulting coefficients were statistically significant and are presented on p. 23 for the interaction effect, and on p. 24 for the main effect on initial profile membership. Additionally, we have revised Tables 4 and 5 to clearly indicate that the reported probabilities are statistically significant.

In short, I think the approach and the model have a lot of potential, but currently there are two many questions, especially concerning quality of measurement, ability of the model to really capture differences in the common and unique components (of alcohol use), and the validity of the results form the applied example to publish this study as is.

We thank the reviewer for these detailed and thoughtful comments. In response, we have added new equations, clarified theoretical assumptions, addressed model estimation procedures, expanded our reporting on data preprocessing and measurement validity, and added transparency about our analytic decisions. We also more fully acknowledge the study's limitations and the need for simulation-based validation.

We appreciate the opportunity to revise our manuscript and thank the reviewers and editor for their constructive feedback. We believe the revisions have substantially strengthened the manuscript. We hope the revised manuscript meets the standards of *Psychology of Addictive Behaviors*, and we look forward to the possibility of having our work considered for publication.

Sincerely,

Ivan Jacob Agaloos Pesigan, Ph.D.
Postdoctoral Fellow, Prevention and Methodology Training Program (PAMT)
Email: ijapesigan@psu.edu