1828
AKADÉMIAI KIADÓ

A clustering based on the dynamics of DSM-5 criteria for gambling disorder: A 5-year follow-up of gamblers with and without gambling disorder

Journal of Behavioral Addictions

15 (2026) 1, 357–370

DOI:
10.1556/2006.2025.00099

GAËLLE CHALLET-BOUJU[1,2*], VINCENT BRAULT[2,3], BASTIEN PERROT[2], JEU-GROUP[†], SOLÈNE DESMÉE[2] and MARIE GRALL-BRONNEC[1,2]

[1] Nantes Université, CHU Nantes, UIC Psychiatrie et Santé Mentale, F-44000 Nantes, France

[2] Nantes Université, Univ Tours, CHU Nantes, INSERM, MethodS in Patients centered Outcomes and HEalth ResEarch, SPHERE, F-44000 Nantes, France

[3] Univ. Grenoble Alpes, CNRS, Grenoble INP, LJK, 38000 Grenoble, France

Received: April 4, 2025 • Revised manuscript received: September 25, 2025 • Accepted: October 30, 2025
Published online: January 12, 2026

FULL-LENGTH REPORT

ABSTRACT

Background and aims: Gambling disorder (GD) follows a non-linear progression over time. Beyond shifts between disordered and non-disordered gambling statuses, the presence of each DSM-5 diagnostic criteria for GD may fluctuate throughout a gambler's trajectory. This study aimed to identify clusters of gamblers based on the evolution of the nine GD diagnostic criteria over time and to characterize these clusters using clinical variables. *Methods:* Data were drawn from the JEU cohort (NCT01207674), a multicenter study of adult gamblers with and without GD, featuring a five-year follow-up. Participants completed structured clinical interviews and self-report questionnaires at baseline and annually. The analysis included only individuals who completed at least one follow-up assessment and met at least one GD diagnostic criterion during the study period ($n = 209$). A mixture model was applied to estimate the probability of meeting each GD criterion over time. *Results:* Four distinct clusters emerged, characterized by a gradient of GD severity and specific criterion-based evolutions. Some criteria fluctuated alongside overall GD severity, while others, particularly "loss of control", persisted even when severity decreased. Additionally, three criteria ("lying", "jeopardization" and "bailout") were more likely to be present only in the two most severe clusters. *Discussion and conclusions:* This study confirms that, beyond the fluctuations of GD states, the presence of individual criteria also varies over time. These findings provide new insights into the dynamic progression of GD and may help tailor therapeutic approaches to better address the specific needs of gamblers at different stages of the disorder.

KEYWORDS

gambling, diagnostic criteria, evolution over time, clustering

[†]Members of the JEU-Group: Marie Grall-Bronnec, Gaëlle Challet-Bouju, Jean-Luc Vénisse, Lucia Romo, Cindy Legauffre, Caroline Dubertret, Jean Adès, Irène Codina, Marc Valleur, Marc Auriacombe, Mélina Fatséas, Jean-Marc Alexandre, Pierre-Michel Llorca, Isabelle Chéreau-Boudet, Christophe Lançon, David Magalon, Amandine Luquiens, Michel Reynaud.

*Corresponding author. UIC Psychiatrie et Santé Mentale, Nantes Université, CHU Nantes, Hôpital Saint-Jacques, 85 Rue Saint Jacques, F-44000 Nantes, France.
Tel.: +33 2 53 48 25 32.
E-mail: gaelle.bouju@chu-nantes.fr

AKJournals

INTRODUCTION

Gambling is a widespread leisure activity. While most people engage in gambling as a recreational activity, some may experience difficulties in regulating their gambling behaviour, which can, over time, lead to the development of an addiction. This creates a spectrum ranging from non-problematic to excessive gambling, including risky behaviors. The most critical point of this spectrum is gambling disorder (GD), a mental disorder classified under addictive disorders in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) (American Psychiatry Association, 2013). GD is defined as persistent and recurrent problematic gambling behavior leading to clinically significant impairment or distress (American Psychiatry Association, 2013). Diagnosis required meeting at least four of the nine DSM-5 criteria, with severity classified as mild (4–5 criteria), moderate (6–7 criteria) or severe (8–9 criteria). GD impacts not only gamblers' physical and mental health but also their social, professional, and personal lives. Financial losses can lead to over-indebtedness or even bankruptcy (Oksanen, Savolainen, Sirola, & Kaakinen, 2018). Individuals with GD often face multiple problems, such as family (conflicts, divorce, isolation), professional (job loss, absenteeism), psychological (anxiety, depression, suicidal thoughts, other addictive behaviors) and physical (fatigue, stress) issues (Francis et al., 2025; Langham et al., 2016; Sean Cowlishaw, Roberts, George, & Kessler, 2025).

DSM-5 criteria are primarily used to diagnose GD and assess its severity. However, two views of GD oppose each other: the first one assumes that individuals with and without GD are qualitatively different, identifiable through a threshold in GD assessments; the other views GD as the extreme end of a severity continuum, resisting clear-cut classification (James, O'Malley, & Tunney, 2014, 2016). Sleczka et al. argue that while the DSM-5 criteria effectively establish GD, a binary approach may oversimplify the syndrome. Analyzing fulfilled criteria (symptoms) as external and observable manifestations of a latent trait (the disorder) can provide deeper insight (Sleczka, Braun, Piontek, Bühringer, & Kraus, 2015). Several studies have examined how specific criteria correlate with different severity levels (S. Cowlishaw et al., 2019; Sacco, Torres, Cunningham-Williams, Woods, & Unick, 2011; Sleczka et al., 2015; Strong & Kahler, 2007; Temcheff, Paskus, Potenza, & Derevensky, 2016; Toce-Gerstein, Gerstein, & Volberg, 2003). Generally, harm-related criteria (jeopardization of important matters and bailout) and withdrawal symptoms were associated with higher severity, whereas other criteria tend to be associated with lower severity (preoccupation with gambling, escapism, tolerance and chasing). Additionally, some researchers questioned whether all criteria hold equal weight in diagnosing GD, suggesting that some may be more indicative of the disorder than others (Lucas et al., 2024; Sleczka et al., 2015).

GD does not always follow a linear and persistent course, and individuals may alternate between symptomatic periods and phases of reduced or absent symptoms (Nower & Blaszczynski, 2008; Pickering, Spoelma, Dawczyk, Gainsbury, & Blaszczynski, 2020; Slutske, 2006; Slutske, Jackson, & Sher, 2003). Indeed, many can regain temporary control through various strategies, such as money management, self-exclusion, or professional treatment. However, beyond shifts in gambling states, the dynamics of how GD criteria evolve over time are not well understood.

We hypothesized that DSM-5 criteria for GD have an unequal and changing weight throughout a gambler's trajectory. The primary objective was to identify clusters of gamblers based on the evolution over time of the nine diagnostic criteria for GD, using a mixture model clustering approach. Then, we sought to describe these clusters in terms of changes in the occurrence of each criterion, GD severity, and treatment initiation. The study's hypotheses and objectives were preregistered in the Open Science Framework (Challet-Bouju, Brault, Desmée, Perrot, & Grall-Bronnec, 2025c).

A better understanding of GD criteria dynamics over time could help identify individuals at risk of developing GD at an earlier stage. Clinically, this could improve treatment by considering not only overall GD severity but also the evolution of specific criteria. Additionally, these findings may contribute to discussions on the relative weight of each criterion in GD diagnosis and potential revisions to current diagnostic criteria.

METHODS

Participants and procedure

Data were drawn from the JEU cohort (NCT01207674), a multicenter study of adult gamblers with or without GD, featuring a five-year follow-up.

Recruitment took place between April 2009 and September 2011, involving clinicians and researchers from seven French institutions. Participants were eligible if they had gambled at least once in the past year and were aged 18 to 65. The only exclusion criteria were an inability to understand French or complete the interviews and questionnaires. To capture a wide range of gambling behaviors and severity levels, recruitment was conducted in care centers (among patients who had initiated treatment within the past six months), gambling venues (casinos, cafés, smoke shops, etc., at the end of a gambling session), and through press advertisements. This approach ensured the inclusion of individuals without GD, individuals with GD not receiving treatment, and individuals with GD currently in care.

The JEU cohort was organized into two phases. The first phase included all participants (n = 628) and consisted of a baseline assessment conducted in a private room at the gambling venue, at a university, or in a hospital, depending on participant preference and venue availability. This assessment included a structured clinical interview and self-report questionnaires. Given the low prevalence of GD in the general population, the sample was intentionally balanced to include approximately equal numbers of gamblers with and without GD. The second phase included only participants who were not in care at baseline (n = 442) and involved a five-year follow-up. These participants were assessed annually using the same measures as in the baseline assessment, adapted for follow-up. For more details on the JEU cohort, please refer to the study protocol (Challet-Bouju et al., 2014).

Selection of participants

The present analysis focused on the dynamics of GD criteria over time, so only data from phase 2 participants were considered. Of the 442 participants in phase 2, we included those who completed at least one follow-up assessment (n = 272) and met at least one GD diagnostic criterion during follow-up (n = 209). This selection ensured the possibility of observing changes in criterion occurrence over time. All available observations for each participant were incorporated into the models, even when some time points were missing. Missing data were not imputed. The participant selection flowchart is presented in Fig. 1.

Measures

The JEU cohort dataset includes a wide range of variables (Challet-Bouju et al., 2014). For this analysis, we selected the most relevant ones.

Sociodemographic characteristics. Data were collected on age, sex, marital status, professional status and education level.

Gambling disorder diagnosis, severity and care. At the time of the study, the DSM-5 had not yet been published in French. Therefore, GD diagnosis was initially assessed using the 10 DSM-IV criteria for pathological gambling: preoccupation with gambling ("preoccupation"), need to gamble with increasing amounts of money to achieve the desired excitement ("tolerance"), repeated unsuccessful attempts to control or stop gambling ("loss of control"), restlessness or irritability when attempting to cut down ("withdrawal"), gambling to escape problems or relieve a dysphoric mood ("escapism"), returning to gambling to recover losses ("chasing"), lying to conceal gambling involvement ("lying"), committing illegal acts to finance gambling ("illegal acts"), jeopardizing or losing significant relationships, jobs, or educational/career opportunities due to

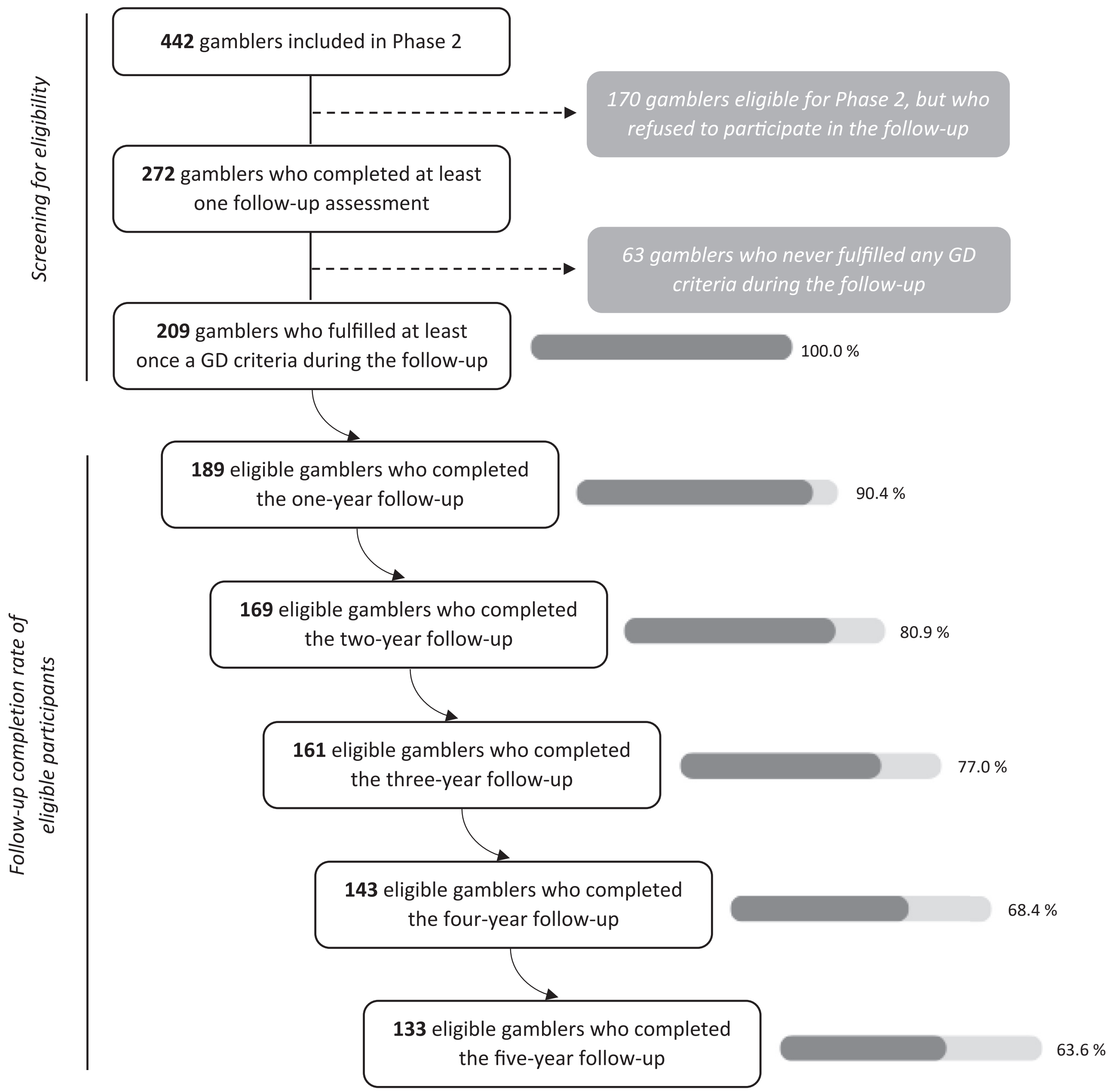


Fig. 1. Flow chart of participants' selection

gambling ("jeopardization"), and relying on others for financial help due to gambling losses ("bailout") (American Psychiatric Association, 2000).

Since changes in the DSM-5 had minimal impact on GD diagnosis (Petry, Blanco, Jin, & Grant, 2014), we retrospectively applied these updates within the JEU cohort: the "illegal acts" criterion was removed, and the diagnostic threshold was lowered to four out of nine criteria (instead of five out of ten in the DSM-IV). GD diagnosis was reassessed at baseline and at each of the five follow-ups, allowing us to track the presence of each DSM-5 criterion over time (baseline, 1-year, 2-year, 3-year, 4-year, and 5-year follow-ups). Additionally, the number of fulfilled criteria at each assessment was used as a measure of GD severity (Strong & Kahler, 2007).

Finally, participants self-reported whether they had received care at each assessment time. Thus, we were able to determine the occurrence of an initiation of care, defined as the first event of care for GD during follow-up.

Statistical analyses

We employed an unsupervised approach to identify clusters of gamblers based on the probability of occurrence of the nine GD diagnostic criteria over time.

First, we recorded the presence or absence of each criterion for each participant at baseline and across five follow-ups. Missing data (e.g., due to missed assessments) were not imputed. To account for variations in gambling trajectories, we defined the starting time (t_0) of the time scale of the model as the age at which each participant declared having gamble for the first time. This ensured that the total duration of gambling experience was considered when analyzing time-related changes in criterion occurrence. Time was treated as a continuous variable, expressed in years.

To classify participants, we applied a logistic mixture model clustering method (McLachlan, Lee, & Rathnayake, 2019), where the variables of interest were the occurrence of each criterion at each assessment. This approach assumes that individuals belong to a heterogeneous population that can be divided into homogenous subgroups (clusters). Mathematically, within each cluster, the nine longitudinal binary observations follow logistic functions linearly influenced by time, with parameters depending on the specific criterion and the cluster (Choi, Dominici, Zeger, & Ouyang, 2005). Parameters of the model and of the clusters were estimated using the Expectation-Maximization algorithm (Dempster, Laird, & Rubin, 1977), while the Bayesian Information Criterion (BIC) was used to determine the optimal number of clusters (Schwarz, 1978). Posterior probabilities were estimated for each participant to ensure the classification was accurate. We also performed a stability analysis to test the robustness of the clustering solution, using a bootstrap resampling strategy. First, to test the sensitivity to sample composition (dependence on specific participants), we performed an individual-level resampling. A total of 1,000 bootstrap samples were generated across ten resampling levels, corresponding to the exclusion of 5–50% of the original participants (100 replicates per level). Second, to test the sensitivity to data incompleteness (i.e., missing visits leading to partial trajectories), we performed an individual-level resampling. A total of 1,000 bootstrap samples were generated across ten resampling levels, corresponding to the exclusion of 5–50% of the observations of participants with more than 2 observations (100 replicates per level). For each replicate (either individual-level or observation-level), the clustering model was re-estimated under the same specifications as the original model (i.e., same number of clusters). The resulting partitions were then compared with the reference solution derived from the full dataset using the Normalized Clustering Evaluation (NCE) criterion (Robert, Vasseur, & Brault, 2021), which quantifies the proportion of consistent classifications between two partitions.

The model estimated the probability of meeting each GD diagnosis criterion over time for each cluster, from gambling initiation onward. Clusters were then characterized based on: (i) the evolution of each criterion's occurrence estimated probability over time, and (ii) the evolution of the number of DSM-5 criteria over time, estimated using the Locally estimated scatterplot smoothing (LOESS) method (Cleveland, Grosse, & Shyu, 1992; Wickham, 2016). LOESS is a nonparametric method that fits a smooth curve through scatterplot data without assuming an underlying structure. Finally, we described the proportion of individuals in each cluster who initiated care during follow-up.

The analysis strategy was preregistered in the Open Science Framework, including the analysis plan (Challet-Bouju, Brault, Desmée, Perrot, & Grall-Bronnec, 2025b) and analysis code (Challet-Bouju, Brault, Desmée, Perrot, & Grall-Bronnec, 2025a).

Ethics

The JEU cohort was approved by the French Research Ethics Committee (CPP) and conducted in accordance with Good Clinical Practice Guidelines and the Declaration of Helsinki. Written informed consent was obtained from all participants before inclusion in the study.

RESULTS

At baseline, the sample had a mean age of 47 years and consisted of 41% women. Nearly half of the participants lived as a couple, while most were active (employed or students) and had at least a high school education. The complete sociodemographic characteristics of the sample at baseline are provided in Table 1.

As displayed in Fig. 1, among the 209 participants included in this study, follow-up completion rates were as follows: 189 (90.4%) at year 1, 169 (80.9%) at year 2, 161 (77.0%) at year 3, 143 (68.4%) at year 4, and 133 (63.6%) at year 5.

Regarding the clustering analysis, Table 2 shows that BIC values decreased from 2- to 4-cluster solutions, and then consistently increased from the 5-cluster solution onward. Thus, we identified the 4-cluster solution as the optimal model. Posterior probabilities were averaged for each cluster

Table 1. Sociodemographic characteristics of the sample at baseline

	N	M (SD) or *n* [%]
Sex (% women)	209	86 [41.1%]
Age (years)	209	46.9 (12.4)
Marital status (% living alone)	209	103 [49.3%]
Educational level (at least a high school education)	208	110 [52.9%]
Professional activity (% active = employed or students)	209	123 [58.9%]

Table 2. Evolution of the BIC criterion according to the number of clusters

K	BIC	Evolution
2	4030.04	
3	3946.48	−83.56
4	3930.58	−15.9
5	3953.79	+23.21
6	3962.98	+9.19
7	4013.16	+50.18
8	4067.68	+54.52
9	4120.17	+52.49
10	4193.68	+73.51
11	4259.89	+66.21
12	4322.1	+62.21
13	4416.66	+94.56
14	4473.65	+56.99

Note: K: number of clusters in each tested solutions.

and are reported graphically in Appendix 1. The medians of posterior probabilities were all very close to 1 (0.97–1.00), with a tight distribution, indicating a high capacity of the model to correctly classify participants in each cluster. Only Cluster 2 shows a wider distribution of posteriori probabilities, but its first quartile was above 0.75, which is still acceptable. The global entropy of the posterior cluster membership probabilities was 37.9. Given that the maximum possible entropy for 209 patients and 4 clusters is 289.7, this corresponds to ∼13% of the maximum. This indicates that most individuals had a clearly dominant cluster assignment, with limited overall uncertainty in the classification.

Results of the stability analyses are displayed in Appendix 2. The results of the individual-level resampling indicated a good stability, with NCE ranging from 0.26 (when 95% of participants were retained) to 0.39 (when 50% were retained). The level of concordance was slightly higher for observation-level resampling, with NCE ranging from 0.22 (when 95% of observations were retained) to 0.39 (when 50% were retained). Concordance increased as a larger proportion of data was retained in both cases, reflecting greater sensitivity under stronger perturbations. Importantly, stability remained well above chance levels even under substantial perturbations, supporting the robustness of the identified clustering solution. These results demonstrate that the four-cluster structure is reproducible across different bootstrap conditions. Each of the four clusters exhibits different dynamics in the probability of meeting the nine GD diagnostic criteria over time. Figure 2 presents the estimated probability of meeting each GD criterion over time for each cluster, from gambling initiation onward. Figure 3 illustrates the smoothed estimate of the number of DSM-5 criteria over time. For both graphical representations, the rolling confidence intervals widen at the early and late time points of the time axis, indicating lower precision of the smoothed estimate. Detailed model specifications (formal definitions and equations) are provided in Appendix 3.

Cluster 1 was the most represented cluster, comprising approximately one-third of the sample (33.5%). The only criteria likely to be present at baseline were “tolerance”, “escapism”, “chasing” and “lying”, with “chasing” being the most probable. Over time, the probability of all the criteria decreased to zero, except for “escapism”, which remained persistent. At baseline, only 6% of gamblers in Cluster 1 met the criteria for GD, all with mild severity; this proportion then dropped to 0% over time. This cluster was characterized by a near-zero number of DSM-5 criteria throughout the study period. No initiation of care was observed in this cluster during follow-up.

Cluster 2 represented 21.1% of the sample. Almost all criteria were likely to be present at baseline, except for “jeopardization” and “bailout”, which remained absent throughout the study. The most probable criteria at baseline were “escapism” (probability = 0.68) and “loss of control” (probability = 0.46). The probability of nearly all criteria present at baseline decreased to between 0 and 0.27 over time, except for “chasing”, which increased significantly, reaching 0.50 at the end of the follow-up period. At baseline, one-quarter of gamblers in Cluster 2 (25.0%) met the criteria for GD, primarily with mild severity (18.2%), and this proportion fluctuated between 7.3 and 13.2% over time. The observed number of DSM-5 criteria remained relatively stable, oscillating between 1 and 2. No initiation of care was observed in this cluster during follow-up.

Cluster 3 accounted for 27.3% of the sample. All criteria were likely to be present at baseline, with “escapism”, “chasing” and “lying” being the most probable (probability = 0.66, 0.72, and 0.77 respectively). While the probability of most criteria decreased over time, the probability of “loss of control” and “escapism” remained stable, whereas “preoccupation” increased. At baseline, a majority of gamblers in Cluster 3 met the criteria for GD (57.9%), primarily with mild (28.1%) or moderate (24.6%) severity. This proportion gradually decreased over time, from 40.7% at the 1-year follow-up to 13.8% at the 5-year follow-up. Initially, the observed number of DSM-5 criteria was close to the threshold for GD diagnosis (4 criteria), and slightly decreased to 2 over time. No initiation of care was observed in this cluster during follow-up.

Cluster 4 was the least represented cluster, comprising 18.2% of the sample. All criteria were likely to be present at baseline, with high probabilities (ranging from 0.67 to 0.98) for most of them, including “tolerance”, “withdrawal”, “loss of control”, “preoccupation”, “escapism”, and “chasing”, with “lying” being somewhat less probable. For half of the criteria (“tolerance”, “withdrawal”, “preoccupation”,

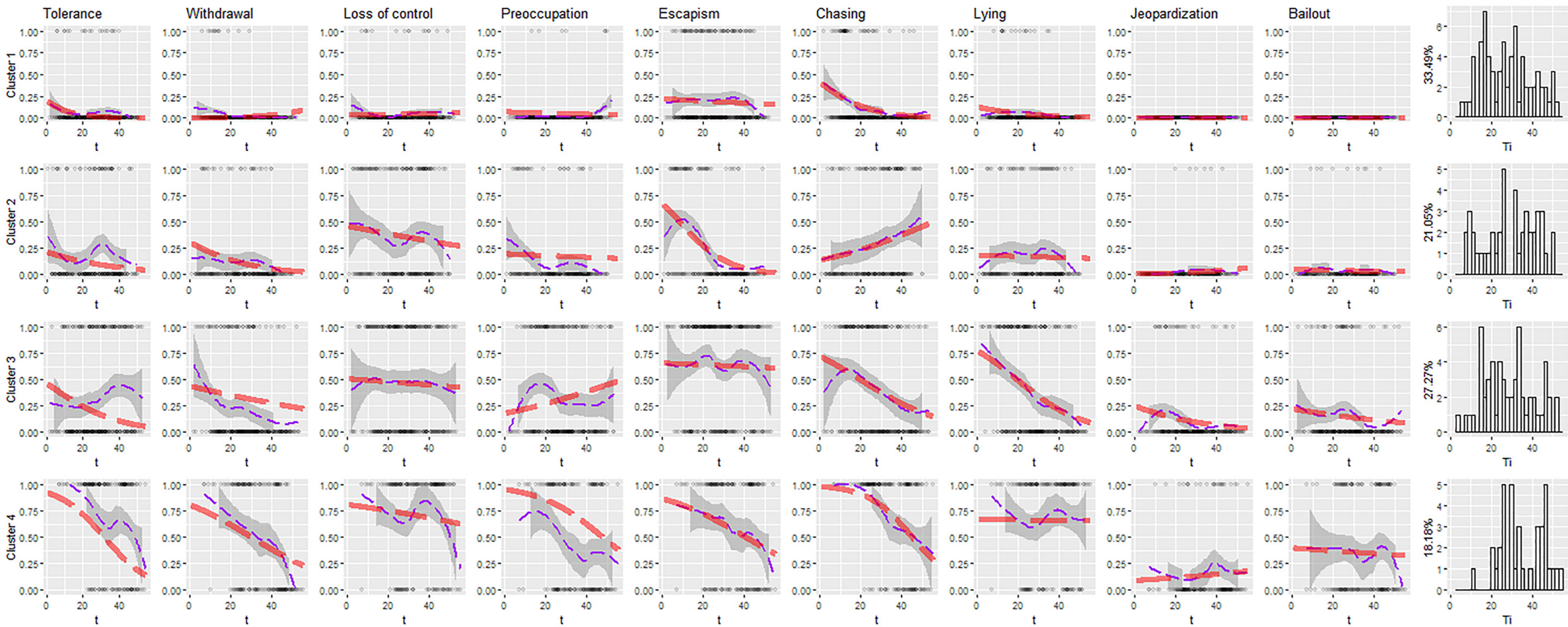


Fig. 2. Scatter plot representing the evolution of the probability of meeting each diagnostic criterion over time (in years) since gambling initiation (in column) estimated by the model, for the 4 identified clusters (in row)

Note: The red dashed line represents the mean probability of fulfilling each DSM-5 criterion for each cluster, as estimated by the mixture model.

The purple dashed line represents a smooth estimate of the raw evolution over time of the proportion of fulfilment of each DSM-5 criterion for each cluster. The gray area represents a rolling 95% confidence interval around this estimate. The black circles represent the true observations, i.e., each circle corresponds to the fact that the individual in the cluster met (probability = 1) or does not met (probability = 0) the criterion at the corresponding moment. Since the variable is binary, points are plotted either at the top ($y = 1$) or bottom ($y = 0$) of the graph.

The x-axis represents the time since gambling initiation, in years. The y-axis represents the probability of meeting each diagnostic criterion.

Histograms on the right-hand side of the figure represent the distribution of time since gambling initiation as reported at baseline, for individuals belonging to each cluster.

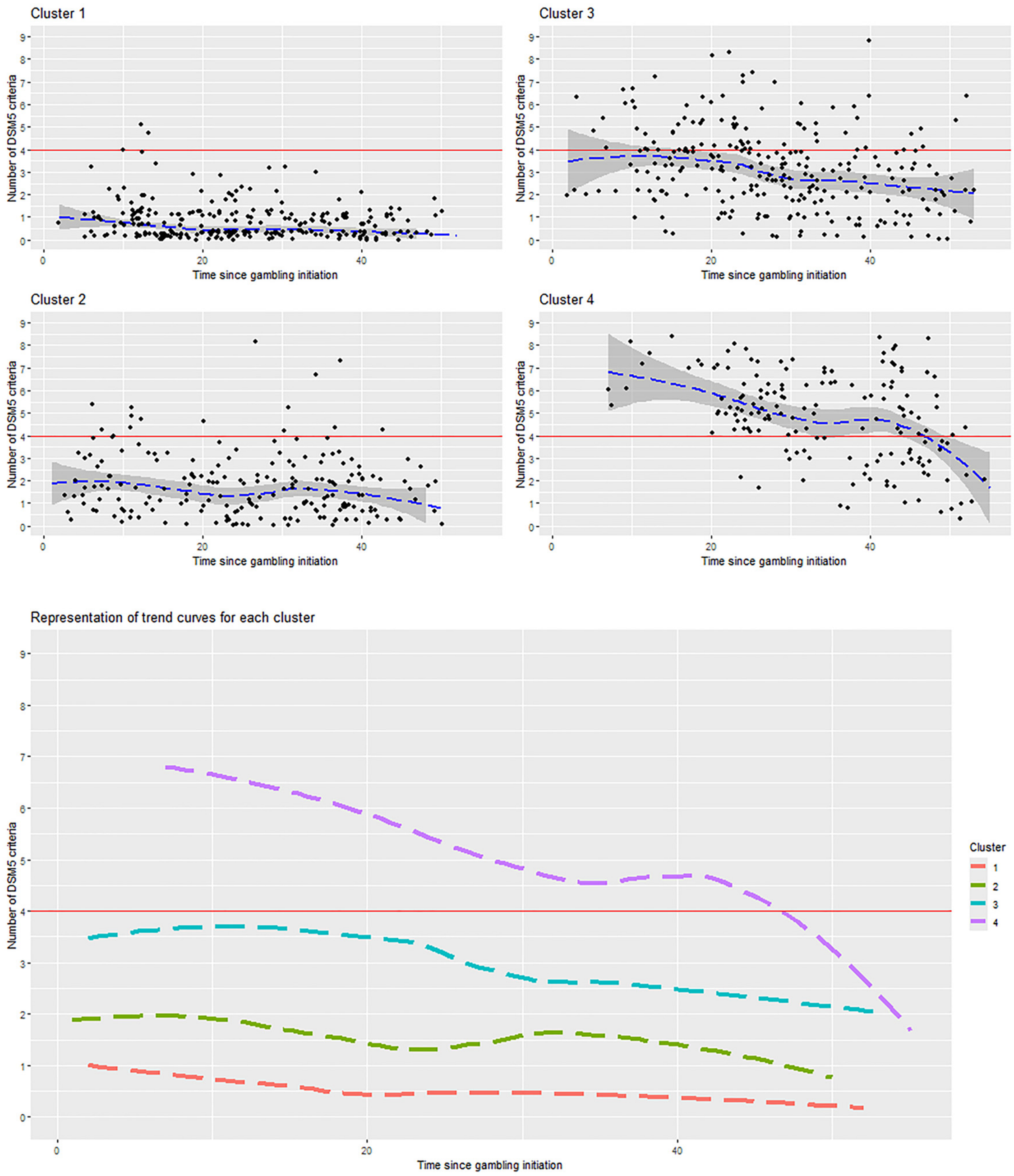


Fig. 3. Scatter plot representing the evolution over time (in years) of the number of DSM-5 criteria, by cluster
Note: Panels 1–4 show the fitted trajectories separately for each cluster. The blue dashed line represents a smooth estimate corresponding to a trend line of the evolution over time of the number of DSM-5 criteria (GD severity), and the gray area represents a rolling 95% confidence interval for this estimate. Each point corresponds to an individual observation; points were displayed with a small random jitter to reduce overlap and improve visibility.
The fifth panel combines all clusters to allow direct comparison of the trajectories.
The red line represents the threshold for GD diagnosis, i.e. 4 fulfilled criteria out of 9.
The *x*-axis represents the time since gambling initiation, in years. The *y*-axis represents the cumulative number of diagnostic criteria met.

"escapism", and "chasing"), probabilities dropped below 0.40 over time. The probabilities of "loss of control" and "lying" remained high and stable throughout follow-up, while the probability of "bailout" remained stable but lower. The probability of "jeopardization" doubled during the follow-up period (from 0.09 to 0.18). Cluster 4 had the highest proportion of gamblers meeting the criteria for GD throughout the study, from 89.5% at baseline to between 87.1% at the 1-year follow-up and 57.9% at the 5-year follow-up. This cluster also exhibited the highest severity levels (26.3% mild, 42.1% moderate, 21.1% severe). The observed number of DSM-5 criteria for GD was high at baseline (around 7 criteria). It initially declined slowly while remaining above the diagnosis threshold in the first years, then sharply dropped below the threshold of 4 by the end of the follow-up period (to approximately 2 criteria). Cluster 4 was the only cluster in which a proportion of gamblers (39%) initiated care for their gambling problem during follow-up.

DISCUSSION

The objective of this study was to identify clusters of gamblers with different dynamics in the probability of occurrence of the nine diagnostic criteria for GD over time. We identified four clusters, each displaying distinct trajectories of GD criteria evolution. These clusters followed a severity gradient from Cluster 1 to Cluster 4 and were also characterized by specific changes in individual criteria over time. Our findings confirm that, beyond fluctuations in GD status, the presence of individual diagnostic criteria also varies over time.

Main findings: a criterion-specific approach

For three of the nine criteria - “withdrawal”, “tolerance” and “preoccupation” -, the probability of occurrence followed the severity gradient between clusters, even as overall disorder severity declines. This suggests that these criteria are particularly relevant for identifying GD and assessing its severity, possibly representing core symptoms of addiction. This finding aligns with a recent network analysis that identified “withdrawal” as the most central criterion for both men and women and ranked “withdrawal”, “tolerance” and “preoccupation” among the four most central criteria based on eigenvector centrality (Lucas et al., 2024). Moreover, the probability of occurrence of these criteria generally decreased over time, mirroring the overall decline in GD severity, which may reflect self-regulation and/or the effects of treatment (in Cluster 4). Indeed, recovery from GD is common (36–39%) among individuals with a history of GD, with natural recovery (i.e., without treatment) accounting for the majority of cases (Slutske, 2006).

The probability of “loss of control” also followed the severity gradient but remained relatively stable over time. This suggests that this symptom tends to persist throughout the gambling trajectory, even when the severity of the disorder diminishes. As Weinstock et al. have pointed out, loss of control is a central characteristic of the disorder (Weinstock, April, & Kallmi, 2017). The study by James et al., which was based on a series of latent class analyses of several prevalent problem gambling assessments in nationally representative British surveys, concluded that loss of control could serve as a marker of the disorder’s severity, thus giving it the status of a qualitative marker (James, O’Malley, & Tunney, 2016). This may support a more categorical rather than dimensional approach to the disorder. Our results are not congruent, as loss of control was observed in clusters 2, 3, and 4, among gamblers with varying degrees of GD severity, which would support the hypothesis of a dimensional approach to the disorder. More than 20 years ago, Toce-Gerstein et al. showed that this criterion could be present in gamblers reporting only one diagnostic criterion, but that its likelihood of occurrence was correlated with the severity of the disorder (Toce-Gerstein et al., 2003). To our knowledge, no previous research has identified the chronicity of this symptom in the gambling course. It is interesting to note that this criterion remains stable even as the severity of the disorder decreases. It is possible to think that loss of control may be a scar of the disorder, persisting even as the gamblers improve or their disorder is resolved. One might assume that they have internalized, from past experiences of failing to control, reduce, or stop gambling behavior, a definitive inability to maintain control.

The other five criteria evolved differently over time, with no clear trend regarding the severity gradient between the clusters and/or specific evolutions over time.

“Escapism” was the only criterion to persist over time in cluster 1, which was the one characterized by the absence of GD over time. Moreover, it did not evolve according to the severity gradient between the clusters, contrary to the other criteria, which may question the relevance of this criterion in the diagnosis of GD. In a study investigating whether escapism has a consistent association with different types of addictive behaviors, the authors found that it was associated with several substance-related addictive behaviors (alcohol consumption, smoking and drug use), but not with gambling (Jouhki & Oksanen, 2022). In the same way, a recent study on excessive online behaviors during the COVID-19 pandemic found that a within-person effect of escapism (changes over time of the level of escapism for a given individual) but not a between-person effect (comparison between individuals with higher *vs* lower levels of escapism) was demonstrated for excessive gambling, contrary to other excessive online behaviors (excessive gaming, excessive internet use) for which both effects were demonstrated (Jouhki, Savolainen, Sirola, & Oksanen, 2022). Thus, the “escapism” criterion may reflect more a motive for the gambling behavior, whatever the severity of gambling-related problems, rather than a symptom of the disorder. It should be considered to remove this criterion for the diagnosis of GD, or at least to give it less weight than the other criteria.

The probability of occurrence of “chasing” decreased in all clusters over time, except for Cluster 2 where it increased constantly. Cluster 2 was characterized by a low level of GD severity over time. Moreover, for all the other clusters, the initial levels of the probability of occurrence of the “chasing” criterion were the highest across all the criteria. This is consistent with previous literature that showed that chasing is a behavior commonly present at subclinical levels of the GD (Toce-Gerstein et al., 2003) and constitute a behavioral marker for GD (Ciccarelli, Cosenza, D’Olimpio, Griffiths, & Nigro, 2019). According to Toce-Gerstein et al., it is the diagnostic criterion most frequently present when only one or two diagnostic criteria are identified, and it is also the most present when the disorder emerges (Toce-Gerstein et al., 2003). It is therefore an essential criterion for marking the transition. Thus, the “chasing” criterion may characterize subclinical or beginning GD. However, this criterion has been identified as the second most central criterion for GD in women (Lucas et al., 2024).

Regarding the last three criteria, the “lying” criterion was the only criterion that persisted at high levels in Cluster 4, the cluster with the highest severity. Moreover, the "jeopardization" and "bailout" criteria were likely to be present only in the two clusters with higher levels of severity (Cluster 3 and Cluster 4), which supports previous literature that found that harm-related criteria are the most associated with higher levels of severity (S. Cowlishaw et al., 2019; Sacco et al., 2011; Sleczka et al., 2015; Strong & Kahler, 2007; Temcheff et al., 2016; Toce-Gerstein et al., 2003). Those three criteria were also the only ones that persisted over time in Cluster 4 despite a decrease in the mean global severity, so they can also represent indicators of gambling problems persistence. Finally, they may be a manifestation of the relationship between the gambler and their close relatives, and more specifically of the loss of trust from the relatives. Indeed, one can imagine that the gambler gradually regains control over gambling, with improvement on GD symptoms, but experiences gambling lapses. The gambler may try to hide the gambling lapses to their relative by lying, which can be discovered by the relatives and lead to the jeopardization of the relationship. And the lapses can also maintain the financial consequences of the GD and necessitate reliance on the partner to get by.

Limitations and strengths

This study had several limitations. First, our sample is not representative of the general population of gamblers (more specifically in terms of prevalence of GD). Indeed, the JEU cohort had a case–control design, with an approximate equality of size between gamblers with and without GD. This choice was performed to be able to observe changes of state in GD status, given that GD has a low prevalence in the general population. However, gamblers included in the JEU cohort were found to have close socioeconomic characteristics with gamblers from the French national prevalence survey (Challet-Bouju et al., 2014). Second, the rate of loss of follow-up was quite high (38%), even if this is expected in such long-term follow-up study. We can hypothesize that those lost to follow-up may be informative, which could have biased the results. Indeed, it is possible that those lost to follow-up did not participate to follow-ups because the severity of their GD increased or on the contrary decreased. Third, we integrated the age at which the gamblers declared having gamble for the first time as the starting time of the model, because each participant was not assessed at the same time of their gambling course. If this approach seemed to us more relevant to take into account the global duration of the gambling course for this time-related model, we cannot exclude two biases linked to this choice. The first is a recall bias, since the age of initiation was self-reported by the participants. Although errors can occur in either direction, we assume that over-reporting (i.e., recalling an older initiation age) would be more likely on average, due to social-desirability pressures to appear as having started later and genuine difficulty recalling the very first gambling episode. This implies underestimation of cumulative exposure (years since initiation) and attenuation of associations with gambling-related outcomes (bias toward the absence of criteria). Therefore, our estimates are likely conservative. The second is a potential confounding generated by the structural correlation between duration of gambling participation and participant age. As such, the observed trajectories may partly reflect age-related patterns rather than pure gambling involvement dynamics. Future research could examine alternative modeling strategies that may help disentangle age-related and trajectory-related patterns. Fourth, the graphical representations displayed in Figs 2 and 3 should be interpreted with caution. Indeed, smoothed trajectories may be less reliable at the boundaries of the time axis, since the smoothing procedure relies on one-sided windows at the earliest and latest time points. This can exaggerate or flatten trends at the margins. Therefore, the shapes of the curves at the beginning and end of the time axis should be interpreted with caution. These limitations are counterbalanced by several strengths of the study. First, GD criteria were explored through face-to-face or telephonic (for follow-ups) interviews based on the DSM, which is much more reliable than self-reported questionnaires. Second, a five-year study with six assessment waves, conducted annually through semi-structured interviews, is very rare in the gambling literature, especially with this sample size (studies with such sample sizes were generally based on telephone-based or online surveys). Third, the inclusion of both gamblers with and without GD is one of the main strengths of our study, because it gave us access to a large range of GD severity and to observe more various gambling trajectories than in studies including only community gamblers or clinical samples. All these strengths have allowed us to analyze for the first time the evolution over time of the GD DSM criteria, while previous literature on this topic was exclusively based on cross-sectional designs.

CONCLUSIONS

To conclude, this is the first study to our knowledge that explored the dynamics of evolution of the GD criteria over time. It highlighted that the number of criteria and the GD status associated do not reflect alone the finer-grained changes observed in the variability of DSM-5 criteria over time. It confirms the assumption that all criteria do not have equal weight in diagnosing GD at different stages of the gamblers' trajectory.

These findings raise the hypothesis that the follow-up of each diagnostic criteria individually may eventually help guide individualized therapeutic approaches. However, further research is needed to validate these subgroups, assess their predictive value, and explore how they could meaningfully inform treatment planning.

Funding sources: This study received funding from two sources. The JEU cohort was supported by both: (i) The joint support of the French Inter-departmental Mission for the fight against drugs and drug addiction (MILDT, currently known as MILDECA) and the French National Institute of Health and Medical Research (INSERM), as part of the call for research projects launched by these two organizations in 2007, for lunching the study, inclusion of participants and baseline assessment; and (ii) A grant from the French Ministry of

Health (PHRC 2009 - RCB 2008-A01188-47), for follow-up assessments. The analyses performed specifically for the present study were funded by the Pari Mutuel Urbain (PMU), a French gambling operator, as part of the implementation of the French obligation to finance scientific studies on gambling and related addictive disorders (Law n° 2010-476 of May 12th modified, art. 3). Nantes University Hospital is the sponsor of this study. There were no constraints on publishing.

Authors' contribution: Study concept and design: GCB, MGB. Obtaining funding: GCB, MGB. Supervision of the JEU cohort: GCB, MGB. Inclusion of participants and collection of data in the JEU cohort: JEU-group; Methodology and statistical analysis of the present study: VB, BP, SD. Interpretation of results: GCB, MGB. Drafting the manuscript: GCB. All authors gave feedback on and approved the final manuscript. All authors had full access to all data in the study, take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflict of interest: GCB and MGB declare that the Endowment Fund of the University Hospital of Nantes received funding from the gambling industry [Française des Jeux (FDJ) and Pari Mutuel Urbain (PMU)] as part of the implementation of the obligation to finance scientific studies on gambling and related addictive disorders (Law n° 2010-476 of May 12th modified, art. 3). This funding is carried out in the form of a sponsorship donated to the Endowment Fund of the University Hospital of Nantes whereas the sponsor of the study is the University Hospital of Nantes, making it possible to guarantee the scientific independence, objectivity and impartiality of the research work. VB, BP and SD declare that they have no competing interests in relation to this work.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders*. Text Revision (DSM-IV-TR) (4th ed.). Washington, DC: American Psychiatric Association.

American Psychiatry Association. (2013). *DSM-5: Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Publishing. (BU Piéron 616.890 75 DSM).

Challet-Bouju, G., Hardouin, J. B., Vénisse, J. L., Romo, L., Valleur, M., Magalon, D., … Grall-Bronnec, M. (2014). Study protocol: The JEU cohort study – Transversal multiaxial evaluation and 5-year follow-up of a cohort of French gamblers. *BMC Psychiatry*, *14*, 226. https://doi.org/10.1186/s12888-014-0226-7

Challet-Bouju, G., Brault, V., Desmée, S., Perrot, B., & Grall-Bronnec, M. (2025a, février 26). Dynamics of DSM-5 criteria for gambling disorder—Analysis code. Consulté 26 février 2025, à l'adresse https://osf.io/kewz6/?view_only=6037d3769c744124b7ac8032cbe8420d.

Challet-Bouju, G., Brault, V., Desmée, S., Perrot, B., & Grall-Bronnec, M. (2025b, février 26). Dynamics of DSM-5 criteria for gambling disorder—Analysis plan. Consulté 26 février 2025, à l'adresse https://osf.io/suwg9/?view_only=27c3029cafb24bb6a225c88c67bd3b39.

Challet-Bouju, G., Brault, V., Desmée, S., Perrot, B., & Grall-Bronnec, M. (2025c, février 26). Dynamics of DSM-5 criteria for gambling disorder—Hypothesis and objectives. Consulté 26 février 2025, à l'adresse https://osf.io/z2ymh/?view_only=49c08e7becbe464aaa7a727c51938d9e.

Choi, L., Dominici, F., Zeger, S. L., & Ouyang, P. (2005). Estimating treatment efficacy over time: A logistic regression model for binary longitudinal outcomes. *Statistics in Medicine*, *24*(18), 2789–2805. https://doi.org/10.1002/sim.2147

Ciccarelli, M., Cosenza, M., D'Olimpio, F., Griffiths, M. D., & Nigro, G. (2019). An experimental investigation of the role of delay discounting and craving in gambling chasing behavior. *Addictive Behaviors*, *93*, 250–256. https://doi.org/10.1016/j.addbeh.2019.02.002

Cleveland, W. S., Grosse, E., & Shyu, W. M. (1992). Local regression models. In *Statistical models in S*. Routledge.

Cowlishaw, S., Merkouris, S. S., Dowling, N. A., Rodda, S., Suomi, A., & Thomas, S. L. (2019). Locating gambling problems across a continuum of severity: Rasch analysis of the Quinte Longitudinal Study (QLS). *Addictive Behaviors*, *92*, 32–37. https://doi.org/10.1016/j.addbeh.2018.12.016

Cowlishaw, S., Roberts, A., George, S., & Kessler, D. (2025). Tackling gambling harm requires a public health approach. *BMJ (Clinical Research Ed.)*, *388*, r447. https://doi.org/10.1136/bmj.r447

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

Francis, J., Fricke, A., Mavranezouli, I., Navein, A., Ritchie, L., Gaskell, M., & Guideline Committee and the Technical Team (2025). Identification, assessment, and management of gambling-related harms: Summary of NICE guideline. *BMJ (Clinical Research Ed.)*, *388*, r323. https://doi.org/10.1136/bmj.r323

James, R. J. E., O'Malley, C., & Tunney, R. J. (2014). On the latent structure of problem gambling: A taxometric analysis. *Addiction (Abingdon, England)*, *109*(10), 1707–1717. https://doi.org/10.1111/add.12648

James, R. J. E., O'Malley, C., & Tunney, R. J. (2016). Loss of control as a discriminating factor between different latent classes of disordered gambling severity. *Journal of Gambling Studies*, *32*(4), 1155–1173. https://doi.org/10.1007/s10899-016-9592-z

Jouhki, H., & Oksanen, A. (2022). To get high or to get out? Examining the link between addictive behaviors and escapism. *Substance Use & Misuse*, *57*(2), 202–211. https://doi.org/10.1080/10826084.2021.2002897

Jouhki, H., Savolainen, I., Sirola, A., & Oksanen, A. (2022). Escapism and excessive online behaviors: A three-wave longitudinal study in Finland during the COVID-19 pandemic. *International Journal of Environmental Research and Public Health*, *19*(19), 12491. https://doi.org/10.3390/ijerph191912491

Langham, E., Thorne, H., Browne, M., Donaldson, P., Rose, J., & Rockloff, M. (2016). Understanding gambling related harm: A proposed definition, conceptual framework, and taxonomy of harms. *BMC Public Health*, *16*(1), 80. https://doi.org/10.1186/s12889-016-2747-0

Lucas, I., Mora-Maltas, B., Granero, R., Demetrovics, Z., Ciudad-Fernández, V., Nigro, G., … Jiménez-Murcia, S. (2024). Network analysis of DSM-5 criteria for gambling disorder: Considering sex differences in a large clinical sample. *European*

Psychiatry: The Journal of the Association of European Psychiatrists, *1–31*. https://doi.org/10.1192/j.eurpsy.2024.22

McLachlan, G. J., Lee, S. X., & Rathnayake, S. I. (2019). Finite mixture models. *Annual Review of Statistics and Its Application*, *6*, 355–378. https://doi.org/10.1146/annurev-statistics-031017-100325

Nower, L., & Blaszczynski, A. (2008). Recovery in pathological gambling: An imprecise concept. *Substance Use & Misuse*, *43*(12–13), 1844–1864. https://doi.org/10.1080/10826080802285810

Oksanen, A., Savolainen, I., Sirola, A., & Kaakinen, M. (2018). Problem gambling and psychological distress: A cross-national perspective on the mediating effect of consumer debt and debt problems among emerging adults. *Harm Reduction Journal*, *15*(1), 45. https://doi.org/10.1186/s12954-018-0251-9

Petry, N. M., Blanco, C., Jin, C., & Grant, B. F. (2014). Concordance between gambling disorder diagnoses in the DSM-IV and DSM-5: Results from the national epidemiological survey of alcohol and related disorders. *Psychology of Addictive Behaviors: Journal of the Society of Psychologists in Addictive Behaviors*, *28*(2), 586–591. https://doi.org/10.1037/a0034661

Pickering, D., Spoelma, M. J., Dawczyk, A., Gainsbury, S. M., & Blaszczynski, A. (2020). What does it mean to recover from a gambling disorder? Perspectives of gambling help service users. *Addiction Research & Theory*, *28*(2), 132–143. https://doi.org/10.1080/16066359.2019.1601178

Robert, V., Vasseur, Y., & Brault, V. (2021). Comparing high-dimensional partitions with the co-clustering adjusted rand index. *Journal of Classification*, *38*(1), 158–186. https://doi.org/10.1007/s00357-020-09379-w

Sacco, P., Torres, L. R., Cunningham-Williams, R. M., Woods, C., & Unick, G. J. (2011). Differential item functioning of pathological gambling criteria: An examination of gender, race/ethnicity, and age. *Journal of Gambling Studies*, *27*(2), 317–330. https://doi.org/10.1007/s10899-010-9209-x

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Sleczka, P., Braun, B., Piontek, D., Bühringer, G., & Kraus, L. (2015). DSM-5 criteria for gambling disorder: Underlying structure and applicability to specific groups of gamblers. *Journal of Behavioral Addictions*, *4*(4), 226–235. https://doi.org/10.1556/2006.4.2015.035

Slutske, W. S. (2006). Natural recovery and treatment-seeking in pathological gambling: Results of two U.S. national surveys. *The American Journal of Psychiatry*, *163*(2), 297–302. https://doi.org/10.1176/appi.ajp.163.2.297

Slutske, W. S., Jackson, K. M., & Sher, K. J. (2003). The natural history of problem gambling from age 18 to 29. *Journal of Abnormal Psychology*, *112*(2), 263–274. https://doi.org/10.1037/0021-843X.112.2.263

Strong, D. R., & Kahler, C. W. (2007). Evaluation of the continuum of gambling problems using the DSM-IV. *Addiction (Abingdon, England)*, *102*(5), 713–721. https://doi.org/10.1111/j.1360-0443.2007.01789.x

Temcheff, C. E., Paskus, T. S., Potenza, M. N., & Derevensky, J. L. (2016). Which diagnostic criteria are most useful in discriminating between social gamblers and individuals with gambling problems? An examination of DSM-IV and DSM-5 criteria. *Journal of Gambling Studies*, *32*(3), 957–968. https://doi.org/10.1007/s10899-015-9591-5

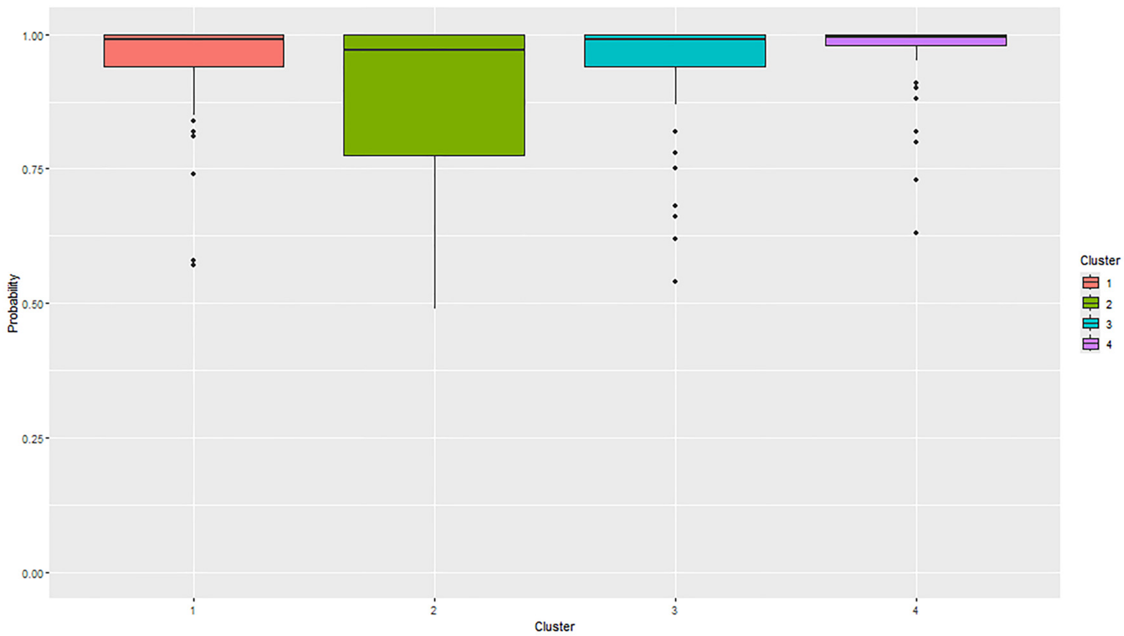
Toce-Gerstein, M., Gerstein, D. R., & Volberg, R. A. (2003). A hierarchy of gambling disorders in the community. *Addiction*, *98*(12), 1661–1672. https://doi.org/10.1111/j.1360-0443.2003.00545.x

Weinstock, J., April, L. M., & Kallmi, S. (2017). Is subclinical gambling really subclinical? *Addictive Behavior*, *73*, 185–191. https://doi.org/10.1016/j.addbeh.2017.05.014

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer Cham. Consulté à l'adresse https://link.springer.com/book/10.1007/978-3-319-24277-4.

Appendices

Appendix 1. Averaged posterior probabilities per clusters



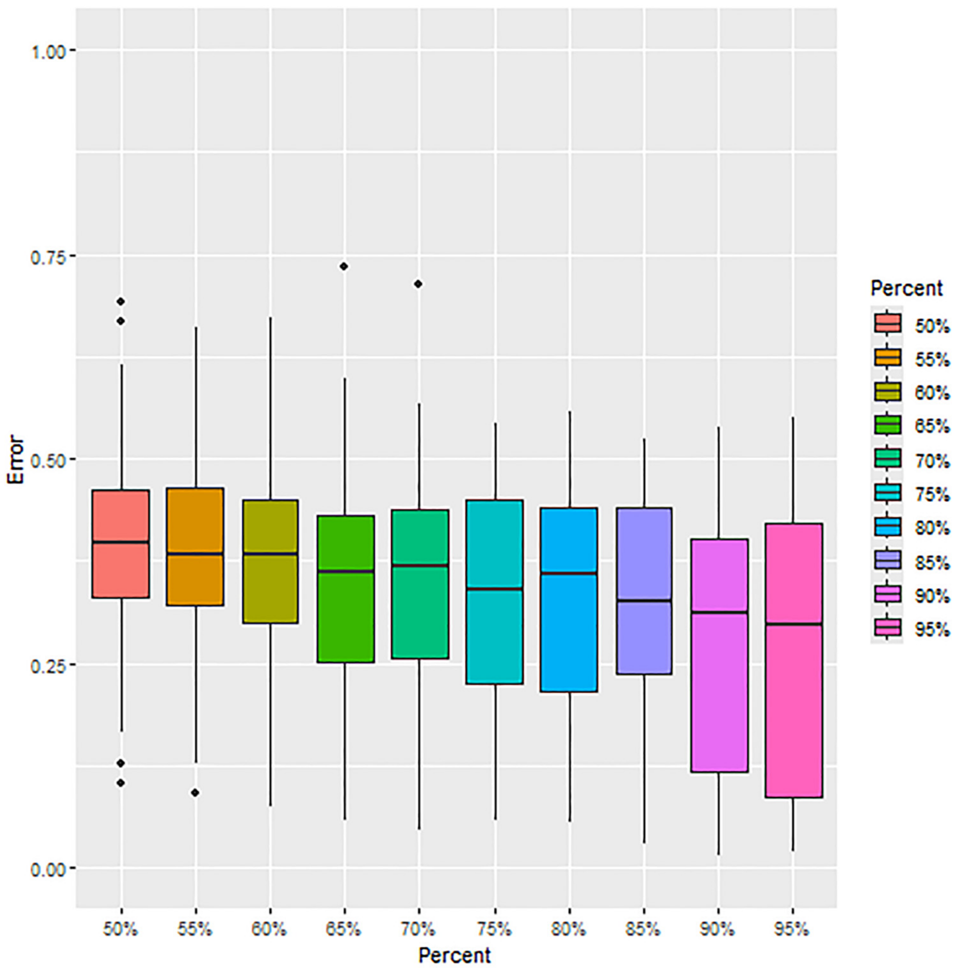
Cluster	Mean	Median	1st quartile	3rd quartile	Minimum	Maximum
1	0.95	0.99	0.94	1.00	0.57	1.00
2	0.87	0.97	0.78	1.00	0.49	1.00
3	0.94	0.99	0.94	1.00	0.54	1.00
4	0.96	1.00	0.98	1.00	0.63	1.00

Note: Average posterior probabilities per cluster were calculated considering that participants were assigned to the cluster in which they had the highest probability of assignment.

Appendix 2. Stability analyses

Individual-level resampling

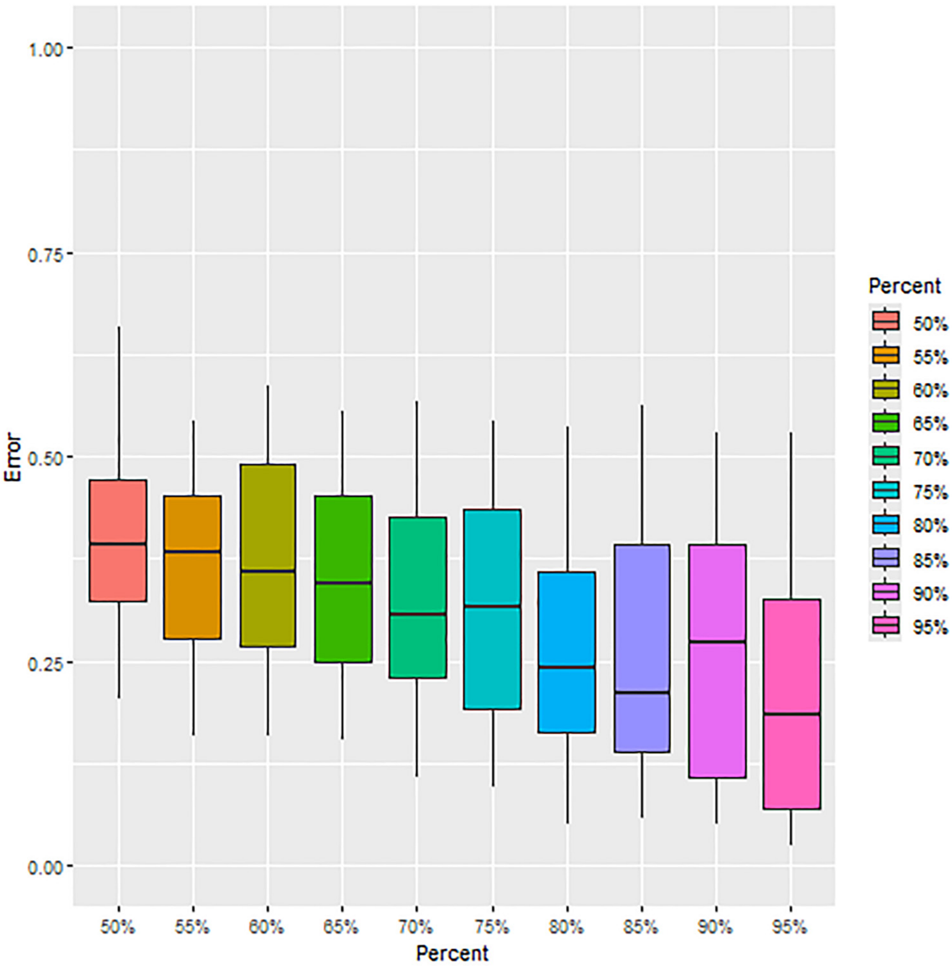
For the individual-level resampling, bootstrap samples were generated by randomly excluding 5–50% of participants from the original dataset (in steps of 5%, 100 replicates per level). This procedure directly tests the sensitivity of the clustering solution to changes in sample composition.



% of the original dataset	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
NCE (sd)	0.390 (0.1)	0.384 (0.1)	0.375 (0.1)	0.350 (0.1)	0.349 (0.1)	0.326 (0.1)	0.331 (0.1)	0.316 (0.1)	0.284 (0.2)	0.259 (0.2)

Observation-level resampling

For the observation-level resampling, we ensured that each participant retained at least two visits, so that longitudinal trajectories remained meaningful. Participants with only two visits (n = 29, 58 visits in total) were left unchanged. Among the remaining 946 visits, observations were randomly removed while maintaining a minimum of two per participant. This allowed us to exclude up to 586 visits in total (corresponding to the situation where all participants are reduced to two visits). The percentages shown in the figure correspond to the proportion of these 586 removable visits that were excluded at each resampling level (e.g., ≈30 visits for 95%, and ≈295 visits for 50%) (also in steps of 5%, 100 replicates per level).



% of the original dataset	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%
NCE (sd)	0.395 (0.1)	0.371 (0.1)	0.371 (0.1)	0.350 (0.1)	0.323 (0.1)	0.313 (0.1)	0.271 (0.1)	0.263 (0.2)	0.259 (0.2)	0.220 (0.2)

Notes: The y-axis represent the value of the NCE (Normalized Clustering Evaluation), and the x-axis the percentage of resampling. The NCE is a normalized version of the classification error. It is used to assess the robustness of a clustering solution by comparing the original partition with partitions obtained under resampling strategies, regardless of the size of clusters. Lower NCE values indicate higher stability of the clustering structure, while higher values reveal greater sensitivity to data perturbations.

Appendix 3. Detailed model specifications

For an individual i in group k, the probability of presenting criterion j at time $\tau_{i.t}$ follows a Bernoulli distribution with parameter $p_{k.j}(\tau_{i.t})$, defined as:

$$p_{k.j}(\tau_{i.t}) = expit\left(a_{k,j} + b_{k,j}\ \tau_{i,t}\right)$$

where $expit$ is the logistic function defined by: $expit(t) = \frac{1}{1+e^{-t}}$

The parameter $a_{k,j}$ corresponds to the probability of having criterion j at gambling onset for group k, and the parameter $b_{k,j}$ represents its temporal evolution.