Applying Data Science to Behavioural Analysis of Online Gambling

Xiaolei Deng, Tilman Lesch & Luke Clark

Centre for Gambling Research at UBC, Department of Psychology, University of British Columbia, Vancouver, Canada.

Address for correspondence: Dr Luke Clark, Centre for Gambling Research at UBC, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, British Columbia, Canada V6T 1Z4.

E-mail: luke.clark@psych.ubc.ca; Xiaolei Deng: xdeng@psych.ubc.ca. Tilman Lesch: Tilman.lesch@psych.ubc.ca

Word counts: abstract 148, main text 2875, 40 references.

This is an Author Accepted Manuscript. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission. The final article is available at: Deng X, Lesch T, Clark L. Applying data science to behavioural analysis of online gambling. *Current Addiction Reports* 2019, in press. doi: 10.1007/s40429-019-00269-9

Abstract

Purpose of review: Gambling operators' capacity to track gamblers in the online environment may enable identification of those users experiencing gambling harm. This review provides an update on research testing behavioural variables against indicators of disordered gambling. We consider the utility of machine learning algorithms in risk prediction, and challenges to be overcome.

Recent findings: Disordered online gambling is associated with a range of behavioural variables, as well as other predictors including demographic and payment-related information. Machine learning is ideally suited to the task of combining these predictors in risk identification, although current research has yielded mixed success. Recent work enhancing the temporal resolution of behavioural analysis to characterize bet-by-bet changes may identify novel predictors of loss chasing.

Conclusions: Data science has considerable potential to identify high-risk online gambling, informed by principles of behavioural analysis. Identification may enable targeting of interventions to users who are most at risk.

Keywords: machine learning, player tracking, loss chasing, self-exclusion, problem gambling

Introduction

The emergence of online gambling and the opportunity to bet via mobile devices has substantially changed the gambling landscape over past two decades. Over the same period, there has been increased recognition of disordered gambling as a form of behavioural addiction and from a public health perspective (1). Public concern about the possible harms of online gambling stems from several factors, including the high availability and immediacy, and the privacy afforded by the online environment. As reviewed by Gainsbury (2) in an earlier issue of *Current Addiction Reports*, empirical support for the specific harmfulness of online gambling is mixed: online gamblers typically engage in a wide range of gambling forms, both online and offline, and this breadth of involvement can be a stronger predictor of gambling problems that online engagement *per se* (3,4).

Online gambling behaviour is inherently linked to single user accounts that can be monitored over time. Such 'behavioural tracking' or 'player tracking' can be used to identify and respond to disordered gambling, presenting an opportunity to create safer gambling environments (5). Advances in data science and the application of these techniques to psychology and psychiatry (e.g. 5,6) are ideally suited to research on online gambling. This reviews provides an update on progress in behavioural analysis of disordered online gambling.

Identifying high-risk gamblers online

Behavioural tracking was pioneered by a group at the Harvard Division on Addictions using a dataset from *bwin.party,* a European platform specializing in sports betting, with the dataset comprising over 48,000 users from 2005 (8). This project has generated over 20 publications, including some independent secondary analyses using the publically available data (www.thetransparencyproject.org) (9–11). Notably, the *bwin* analyses aggregated the gambling behavioural data by day (e.g. total bets per day, number of active betting days). Consequently, these studies offer limited resolution *within* a session of play (see Bet by Bet Behaviour, below).

A basic obstacle in identifying high-risk users is the external verification of which users have gambling problems. In the absence of clinical diagnoses on a vast dataset, researchers must rely on proxy indicators. Braverman & Shaffer (12) focused on 530 *bwin* users who closed their accounts, some of whom cited gambling problems as their reason for account closure. A related approach uses voluntary self-exclusion (VSE) (13–16), a common responsible gambling tool in which the gambler enters a contract with the operator to block his or her access for a set period of time. Around three-quarters of VSE enrollees tend to meet criteria for problem gambling (17,18). This supports the utility of VSE as an indicator - although many individuals with gambling problems do not enrol in VSE programs, and it is possible that self-excluders may have discrete psychological characteristics (e.g. greater insight into their disordered behaviour). Later *bwin* studies combined a number of 'red flag' indicators: account closure, complaints to the website, or requests to increase spending limits (19,20). A subset of *bwin* users also completed the 3-item Brief Biosocial Gambling Screen, a screening instrument that was hosted on the website (4,19). Although such screening is perhaps the most definitive procedure, only 2.2% of users submitted BBGS scores, raising concerns about their representativeness.

The *bwin* analyses and subsequent research using other datasets have shown that problem gambling indicators are associated with a number of behavioural variables that can be derived from gambling play. In Table 1, a simple way to classify these measures is as *monetary* markers, such as net loss or average bet size, and *non-monetary* markers such as session length. Braverman & Shaffer (12) distinguished four groups of *bwin* variables: i) gambling *frequency* (e.g. active betting days), ii) gambling *intensity* (e.g. total bets, average bet size), iii) gambling *variability*; for example, more variation (standard deviation) of the average bet from session to session (12,13,19), iv) gambling *trajectory.* With regard to trajectory, an increasing wager size in the first month after registering an account predicted later account closure (12) and similar increases were seen in the immediate days leading up to account closure (22). It is also recognized that users with likely gambling problems tend to gamble across a greater number of gambling games on the website. This can be considered a distinct form of variability, termed breadth of involvement (4,9).

Several studies have tested multiple variables in predicting problem gambling. For example, Gray et al (20) compared *bwin* gamblers with (n = 2,042) and without (n = 2,014) the red flag indicators on 27 behavioural variables using discriminant function analysis. Non-monetary variables reflecting intensity (active betting days, bets per day) best distinguished the two groups, particularly using game-specific measures from live action sports. In the GTECH dataset (13), gamblers with a record of self-exclusion displayed higher gambling losses and loss variability (i.e. monetary variables), but did not differ on some other measures including bets per day or the number of different games played. Other analyses have used clustering techniques to identify latent subgroups of gamblers who may be high risk, again pointing to high wager variability as a useful marker (12,19).

Machine Learning

Recent analyses have begun to apply machine learning to the challenge of predicting high-risk gamblers. This collection of techniques has been widely embraced in psychology and psychiatry; for example, in multivariate prediction of psychiatric diagnoses (23) or selecting medications (6). A recent study in pathological gamblers sought to classify cases vs healthy controls using the Big Five personality variables, with 77% overall accuracy (22*). In applying machine learning to online behavioural tracking data, the analysis is presented with a number of 'input variables' (e.g. Table 1) and is informed about each subject's status as belonging to a 'target group' (i.e. with an indicator for problem gambling) or a control group. Compared to classical statistics, machine learning is expressly designed for predictive modelling, and captures complex interactions between predictors by default. Model performance is indicated by accuracy, the percentage of subjects who are correctly classified by the model, although this metric is sensitive to unbalanced datasets; for example, if a dataset comprises 75% target cases, the chance level of accuracy is 75%. Other performance measures are better suited to unbalanced data, including the Area Under the Curve of the Receiver Operating Characteristic (AUROC), which characterizes the balance of sensitivity and specificity over a range of discrimination thresholds.

Using the bwin dataset, Philander (11) tested a number of data mining procedures in classifying account closers who did (targets n = 176) vs did not (controls n = 354) specify gambling problems as their reason for account closure. Input variables were behavioural markers of frequency, intensity, variability and trajectory, as well as basic demographics (age, gender, country of residence). Philander trained the model on 70% of available data and then tested the resulting algorithm on the remaining 30%, termed a 'hold out' procedure. The best-performing machine learning model was a random forest procedure: this achieved close to 100% accuracy on the training dataset, but test accuracy was markedly lower (0.66 - 0.67) and the crucial AUROC measures were barely above chance (0.50 – 0.55). In considering this modest performance, it is important to recognize that the control group were individuals that had nevertheless closed their accounts. In a large sample of online poker players who had enrolled in self-exclusion (n = 1966), Luquiens et al (16*) found that the stated reason for self-exclusion (gambling problems versus 'commercial reasons') offered limited discrimination in relation to gambling behaviour, in three machine learning models (highest AUROC 0.57). In the poker players, 68% returned to the website to gamble once their self-exclusion expired, and 60% of those individuals went on to self-exclude again.

A study using the GTECH data presents a more optimistic picture for the use of machine learning. Percy et al (14*) employed self-exclusion as the indicator for problem gambling (n = 176 vs 669 non-VSE controls), with 33 input variables reflecting frequency, intensity, variability, trajectory, along with the basic demographic variables. A random forest procedure showed the highest classification performance, with AUROC of 79% (and accuracy 80%) in the main analysis. Another study (15*) using a newer (2015) dataset from *bwin* incorporated payment-related variables (e.g. deposits and withdrawals from the account) for the first time. This balanced dataset comprised 1,348 users with and without self-exclusion records. Using artificial neural networks, the best-performing model achieved an accuracy of 72%. Other work has examined input variables based on the text analysis of email correspondence with the gambling operator (150 VSE vs 150 controls), which achieved classification accuracy of 78% (25).

Given the limited number of available datasets, it is hard to know to what extent these differences in predictive performance reflect meaningful decisions around the choice of input variables, or simple idiosyncrasies of the datasets, such as thresholds for data inclusion. There is a pressing need for additional datasets to establish generalizability. Notably, Percy et al. included trajectory variables relative to a baseline period for each user. The trajectory information could account for the superior performance, but analysis of trajectories entails decisions regarding the amount of baseline data for prediction, and how to identify users for whom gambling is already disordered at baseline. In Philander and Percy et al, the unbalanced datasets complicate the interpretation of model performance, although real-world data is likely to be unbalanced given the relatively low base-rate of problem gambling (25). Taking these studies together, a conservative conclusion is that *combining* these various tiers of information (betting behaviour, payment behaviour, correspondence) could lead to meaningful levels of predictive performance.

Bet by Bet Behaviour

The research to date that has sought to identify high-risk gamblers from behavioural tracking data has relied on extensive data aggregation. More fine-grained analyses of bet-by-bet behaviour *within* a session of play may be able to capture loss chasing, the tendency to continue gambling or increase one's bet in an effort to recover debts. Chasing is often regarded as a defining feature of problem gambling (26,27), and it is one of the few criteria of problem gambling that may be detectable behaviourally from player tracking data.

Two studies on the *bwin* dataset looked at behavioural dynamics relevant to chasing over a longer timescale of days or weeks (see Table 1). Adami et al (9) hypothesized that unsustainable gambling could be expressed in a 'sawtooth' pattern of wagering over successive days, whereby an initial ramping up of bet amount is followed by a rapid crash as funds are exhausted. There was some support for this pattern based on a clustering analysis. In Ma et al (10), the *bwin* data were aggregated on a weekly basis across 8 months of play. Using the weekly total wager as the dependent variable, gains and losses from the prior week were seen to have opposing effects. Specifically, recent losses reduced the amount gambled in the subsequent week. At the same time, *cumulative* gains and losses (i.e. summed across the period until the current week) both predicted increased bet amount, consistently with a longer term chasing influence. This paper is important for showing that the impact of prior losses changes (and in fact reverses) differs according to whether the losses were recent or chronic.

The first investigation of more detailed bet-by-bet behaviour analyzed 600,000 hands (n = 2,678 users) from the Full Tilt online poker site, to examine how players responded following high-magnitude gains and losses (28). The study focused on high stake games with 25$/50$ blinds, and monitored betting behaviour over the 12 hands that followed gains or losses greater than $1000. Following major losses, players displayed looser betting strategies, consistent with a loss chasing response. Players became more aggressive in their play after major wins.

Other analyses of bet-by-bet behaviour have focussed on specific psychological phenomena. An analysis of 776 online sports bettors (29) examined how winning or losing streaks affected the success of subsequent bets. In a thought-provoking finding, gamblers were more likely to lose after losing streaks, seemingly as a result of accepting riskier bets. Conversely, gamblers were more likely to win after winning streaks, seemingly from placing safer bets. Xu & Harvey interpret their data as substantiating both the hot hand belief and gambler's fallacy. Leino and colleagues (30*) looked at land-based gambling terminals in Norway, where such machines can only be accessed via a personal ID card that enables behavioural tracking (as well as state-mandated loss limits). Their analysis tested the likelihood of continuing a gambling session as function of different bet outcomes, focussing on 'losses disguised as wins', where small wins can be less than the wager placed on a given bet (31). With losses disguised as wins as the reference category, full losses significantly reduced the likelihood of continued gambling, whereas full wins significantly increased continued play. Their finding that losses tended to reduce persistence at an overall level is clearly pertinent to loss chasing, but it is nevertheless possible that opposing results (increased continuation after losses) could have been present in a subset of gamblers such as those with gambling problems.

Overall these studies show considerable promise in identifying behavioural markers of loss chasing using within-session metrics, but existing studies are yet to link these behavioural expressions with red flag indicators of disordered gambling. Given the overwhelming richness of these datasets, it is perhaps unsurprising that studies to date have been narrow in scope and highly hypothesis-driven. Moreover, in focussing on the *typical* (i.e. average) response to losing streaks, researchers may overlook of individual 'outliers' who display an atypical pattern of behaviour that could be useful in risk identification.

Conclusions

Online gambling datasets offer a valuable and rich resource for understanding problematic gambling, and guiding interventions to reduce harm in the online environment. Contrasting with most behavioural research on gambling, these studies consider gamblers who are using their own funds to play games of their own choosing, thus gambling in an entirely naturalistic manner. It is evident that disordered gambling in such users is expressed through multiple behavioural markers, including both monetary and non-monetary variables, as well as variables reflecting frequency, intensity, variability, and trajectory (see Table 1). Researchers are increasingly able to probe behavioural tracking data at higher resolutions, including bet-by-bet analyses. A challenge moving forward is how to integrate new insights from data science (i.e. from de-identified 'big data') with existing psychological knowledge on the detailed characterization of traits and cognitive processes in disordered gambling; for example do predisposing traits such as impulsivity specifically predict behavioural markers such as bet escalation or breadth of involvement?

Machine learning is ideally suited to the task of differentiating behavioural patterns in those with disordered gambling from healthy gamblers. These algorithms can then be used to search for ‘at risk’ individuals in new data, based on their behavioural similarities to the target group (e.g. self-excluders). Nonetheless, current research is dominated by small number of datasets, particularly in relation to North American online gambling. It should be recognized that the widely-used indicators of problem gambling (e.g. self-exclusion, account closure) are imperfect, and algorithms will need to generalize beyond such specific indicators. Behavioural tracking data are typically specific to one operator, and the extent of users’ engagement on other platforms (or offline) is unknown. Lastly, concerns exist around the lack of transparency of machine learning as a multivariate ‘black box’ procedure, although solutions to explaining individual predictions are becoming available (32).

We take care to distinguish the use of data analytics in identification of high-risk online gambling from research testing interventions that might be directed at such users. Gambling operators may offer ‘Responsible Gambling’ intervention tools to *all* users, or may direct such interventions to a subset of users. This targeting may harness machine learning algorithms but could alternatively be based on simple risk factors (e.g. male gender) that do not require sophisticated algorithms. For the gambling operator, there is then a spectrum of possible actions, with enforced account closure at one extreme. At intermediate levels, users could receive in-game messages (e.g. showing time on device) (33), personalized expenditure feedback (e.g. when depositing funds) (34,35), communication of their high-risk status and help-seeking resources (36), or be incentivized to engage with limit-setting tools (37). Current research on the efficacy of these interventions is mostly based on users who ‘opt in’ to such programs, with only a small number of studies using randomized designs that afford stronger conclusions (34). There is some evidence that gamblers have positive attitudes towards such tools (38), and even to more extreme systems that enforce mandatory monthly loss limits (39), but some research also questions whether some popular tools are actually effective, e.g. limit-setting prompts (40*,41).

Although the present paper has focussed on behavioural markers, it should be recognized that online gambling operators have access to additional measures that may be sensitive to disordered behaviour, including payment behaviour (15), time of day (e.g. night-time gambling) (42) and users' correspondence with the website via email or chatrooms (25). A further tier of information may derive from behavioural interactions with a website's Responsible Gambling resources; it is likely that many gamblers may investigate such materials some time before enacting these plans (e.g. self-exclusion enrollment). Finally, how and when should risk scores be presented to maximise their impact and intended consequences? In using behavioural markers to direct interventions at high-risk gamblers, the field of data analytics must be considered alongside research on risk communication (43) and human computer interaction (34).

Author Note

Luke Clark is the Director of the Centre for Gambling Research at UBC, which is supported by funding from the Province of British Columbia and the British Columbia Lottery Corporation (BCLC), a Canadian Crown Corporation. The BCLC and BC Government no constraints on publishing. LC has received speaking or reviewing honoraria from Svenska Spel (Sweden), National Association for Gambling Studies (Australia), National Center for Responsible Gaming (US), and Gambling Research Exchange Ontario (Canada). He has not received any further direct or indirect payments from the gambling industry or groups substantially funded by gambling. He has received royalties from Cambridge Cognition Ltd. relating to neurocognitive testing. XD and TL report no conflicts of interest.

This work was supported by the Centre for Gambling Research at UBC core funding from the Province of British Columbia and the British Columbia Lottery Corporation (BCLC), as well as a research grant from BC Ministry of Finance to LC. XD is supported by a 4 yr fellowship funding from UBC. LC receives funding from the Natural Sciences and Engineering Research Council (Canada).

Compliance with Ethics Guidelines

Conflict of Interest

Dr. Clark reports grants from Centre for Gambling Research at UBC, grants from BC Ministry of Finance, during the conduct of the study; grants from Natural Sciences and Engineering Research Council of Canada, personal fees from Speaker honorarium, Svenska Spel (Sweden),

personal fees from Speaker honorarium, National Association for Gambling Studies (Australia), personal fees from Reviewer honorarium, National Center for Responsible Gaming (US), personal fees from Reviewer honorarium, Gambling Research Exchange Ontario (Canada)., outside the submitted work; and LC is the Director of the Centre for Gambling Research at UBC, which is supported by funding from the Province of British Columbia and the British Columbia Lottery Corporation (BCLC), a Canadian Crown Corporation. The BCLC and BC Government impose no constraints on publishing.

Human and Animal rights and Informed Consent

This article does not contain any studies with human or animal subjects performed by any of the authors

References

Annotations:

Dwyer et al 2018: Excellent primer on the history and application of machine learning in psychiatry and clinical psychology, including methodological decisions such as feature selection and cross-validation.

Percy et al 2016: Machine learning analysis of self-exclusion in the European GTECH dataset. Compared multiple techniques including logistic regression. A random forest model achieved highest performance (AUROC = 79%) in predicting problematic gambling.

Haeusler 2016: The first study to consider online financial behaviours (e.g. amount and number of deposits and withdrawals) in predicting self exclusion. Used artificial neural networks as a form of machine learning to show that payment behaviours achieve a classfication rate of 72%.

Cerasa et al 2018: One of the first studies to apply machine learning to classifying pathological gamblers versus healthy controls, showing 77% overall accuracy using the Big Five personality variables.

Luquiens et al 2018: Reports behavioural tracking over a 6 year period of 1,996 online poker players who self-excluded from the winamax platform. Four machine learning models displayed modest performance differentiated gamblers by their stated reason for self-exclusion (gambling problems versus commercial reasons).

Leino et al 2016: A Norwegian study looking at trial-by-trial behaviour following 'Losses Disguised as Wins' in land-based electronic gaming machines (EGMs). LDWs increased the likelihood of continuing betting compared to full losses.

Wohl et al 2017: Intervention study in casino gamblers playing on a loyalty card, who received personalized expenditure feedback. Those gamblers who underestimated their losses showed reduced gambling over 3 month monitoring.

Ivanova et al 2019: Randomized controlled trial in online slots gamblers, comparing gambling losses in groups who received a limit-setting prompt upon registration, before or after their first deposit, or no prompt (>1000 per group). Prompted groups were more likely to set limits but did not differ in subsequent losses over 90 day follow-up.

1. Clark L. Disordered gambling: the evolving concept of behavioral addiction. Ann N Y Acad Sci. 2014; 1327(1):46–61.

2. Gainsbury SM. Online gambling addiction: the relationship between internet gambling and disordered gambling. Curr Addict Reports 2015; 2:185–93.

3. Philander KS, MacKay T-L. Online gambling participation and problem gambling severity: is there a causal relationship? Int Gambl Stud 2014; 14: 214–27.

4. LaPlante DA, Nelson SE, Gray HM. Breadth and depth involvement: Understanding Internet gambling involvement and its relationship to gambling problems. Psychol Addict Behav 2014; 28(2):396–403.

5. Griffiths MD. Internet gambling, player protection, and social responsibility. In: Williams RJ, Wood RT, Parke J, editors. The Routledge International Handbook of Internet Gambling. London: Routledge; 2012. p. 227–49.

6. Chekroud AM, Zotti RJ, Shehzad Z, Gueorguieva R, Johnson MK, Trivedi MH, et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. Lancet Psychiatry. 2016;3(3):243–50.

7. Dwyer DB, Falkai P, Koutsouleris N. Machine Learning Approaches for Clinical Psychology and Psychiatry. Annu Rev Clin Psychol 2018;14(1): 91–118.

8. Shaffer HJ, Peller AJ, LaPlante DA, Nelson SE, LaBrie RA. Toward a paradigm shift in Internet gambling research: From opinion and self-report to actual behavior. Addict Res Theory. 2010;18(3):270–83.

9. Adami N, Benini S, Boschetti A, Canini L, Maione F, Temporin M. Markers of unsustainable gambling for early detection of at-risk online gamblers. Int Gambl Stud 2013;13(2):188–204.

10. Ma X, Kim SH, Kim SS. Online gambling behavior: the impacts of cumulative outcomes, recent outcomes, and prior use. Inf Syst Res. 2014;25:511–27.

11. Philander KS. Identifying high-risk online gamblers: a comparison of data mining procedures. Int Gambl Stud 2014;14(1):53–63.

12. Braverman J, Shaffer HJ. How do gamblers start gambling: identifying behavioural markers for high-risk internet gambling. Eur J Public Health 2012; 22(2):273–8.

13. Dragicevic S, Percy C, Kudic A, Parke J. A descriptive analysis of demographic and behavioral data from Internet gamblers and those who self-exclude from online gambling platforms. J Gambl Stud. 2015;31(1):105–32.

14. Percy C, França M, Dragičević S, d'Avila Garcez A. Predicting online gambling self-exclusion: an analysis of the performance of supervised machine learning models. Int Gambl Stud. 2016;16(2):193–210.

15. Haeusler J. Follow the money: using payment behaviour as predictor for future self-exclusion. Int Gambl Stud. 2016;16(2):246–62.

16. Luquiens A, Vendryes D, Aubin HJ, Benyamina A, Gaiffas S, Bacry E. Description and assessment of trustability of motives for self-exclusion reported by online poker gamblers in a cohort using account-based gambling data. BMJ Open. 2018;8(12):1–8.

17. Hayer T, Meyer G. Internet self-exclusion: characteristics of self-excluded gamblers and preliminary evidence for its effectiveness. Int J Ment Health Addict. 2011;9(3):296–307.

18. McCormick A V., Cohen IM, Davies G. Differential effects of formal and informal gambling on symptoms of problem gambling during voluntary self-exclusion. J Gambl Stud 2018; 34: 1013-1031.

19. Braverman J, LaPlante DA, Nelson SE, Shaffer HJ. Using cross-game behavioral markers for early identification of high-risk internet gamblers. Psychol Addict Behav 2013;27(3):868–77.

20. Gray HM, LaPlante DA, Shaffer HJ. Behavioral characteristics of Internet gamblers who trigger corporate responsible gambling interventions. Psychol Addict Behav 2012; 26(3):527–35.

21. Tom MA, LaPlante DA, Shaffer HJ. Does Pareto Rule Internet Gambling? Problems Among The "Vital Few" & "Trivial Many." J Gambl Bus Econ 2014;8(1):73–100.

22. Xuan Z, Shaffer H. How do gamblers end gambling: longitudinal analysis of Internet gambling behaviors prior to account closure due to gambling related problems. J Gambl Stud 2009;25(2):239–52.

23. Ahn WY, Vassileva J. Machine-learning identifies substance-specific behavioral markers for opiate and stimulant dependence. Drug Alcohol Depend 2016;161:247–57.

24. Cerasa A, Lofaro D, Cavedini P, Martino I, Bruni A, Sarica A, et al. Personality biomarkers of pathological gambling: a machine learning study. J Neurosci Methods 2018;294:7–14.

25. Haefeli J, Lischer S, Haeusler J. Communications-based early detection of gambling-related problems in online gambling. Int Gambl Stud 2015;15(1):23–38.

26. Gainsbury SM, Russell A, Hing N, Wood R, Blaszczynski A. The impact of internet gambling on gambling problems: a comparison of moderate-risk and problem Internet and non-Internet gamblers. Psychol Addict Behav. 2013;27(4):1092–101.

27. Temcheff CE, Paskus TS, Potenza MN, Derevensky JL. 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. J Gambl Stud 2016;32:957–68.

28. Smith G, Levere M, Kurtzman R. Poker player behavior after big wins and big losses . Manage Sci. 2009;55:1547–55.

29. Xu J, Harvey N. Carry on winning: The gamblers' fallacy creates hot hand effects in online gambling. Cognition 2014;131(2):173–80.

30. Leino T, Torsheim T, Pallesen S, Blaszczynski A, Sagoe D, Molde H. An empirical real-world study of losses disguised as wins in electronic gaming machines. Int Gambl Stud 2016;16(3):470–80.

31. Dixon MJ, Harrigan KA, Sandhu R, Collins K, Fugelsang JA. Losses disguised as wins in modern multi-line video slot machines. Addiction 2010;105:1819–24.

32. Chekroud AM, Foster D, Zheutlin AB, Gerhard DM, Roy B, Koutsouleris N, et al. Predicting barriers to treatment for depression in a U.S. national sample: a cross-sectional, proof-of-concept study. Psychiatr Serv 2018;69(9):927–34.

33. Auer M, Malischnig D, Griffiths M. Is "pop-up" messaging in online slot machine gambling effective as a responsible gambling strategy? J Gambl Issues. 2014;(29):1–10.

34. Auer MM, Griffiths MD. Personalized behavioral feedback for online gamblers: A real world empirical study. Front Psychol. 2016;7(NOV):1–13.

35. Wohl MJA, Davis CG, Hollingshead SJ. How much have you won or lost? Personalized behavioral feedback about gambling expenditures regulates play. Comput Human Behav. 2017;70:437–45.

36. Wood RTA, Wohl MJA. Assessing the effectiveness of a responsible gambling behavioural feedback tool for reducing the gambling expenditure of at-risk players. Int Gambl Stud 2015;15(2):1–16.

37. Wohl MJA. Loyalty programmes in the gambling industry: potentials for harm and possibilities for harm-minimization. Int Gambl Stud 2018;18(3):495–511.

38. Forsstrom D, Jansson-Frijmark M, Hesser H, Carlbring P. Experiences of Playscan: Interviews with users of a responsible gambling tool. Internet Interv 2017;8:53–62.

39. Auer M, Reiestad SH, Griffiths MD. Global limit setting as a responsible gambling tool:

what do players think? Int J Ment Health Addict. 2018;1–13.

40. Ivanova E, Magnusson K, Carlbring P. Deposit limit prompt in online gambling for reducing gambling intensity: A randomized controlled trial. Front Psychol. 2019;10:1–11.

41. Hollingshead SJ, Wohl MJA, Santesso D. Do you read me? Including personalized behavioral feedback in pop-up messages does not enhance limit adherence among gamblers. Comput Human Behav. 2019;94:122–30.

42. PricewaterhouseCoopers. Remote Gambling Research Interim report on Phase II [Internet]. 2017. Available from: https://about.gambleaware.org/media/1549/gamble-aware_remote-gambling-research_phase-2_pwc-report_august-2017-final.pdf

43. Rakow T, Heard CL, Newell BR. Meeting three challenges in risk communication. Policy Insights from Behav Brain Sci 2015;2(1):147–56.

Table 1. Behavioural markers that can be derived from online gambling play, and their linkages with indicators for disordered gambling

Behavioural Marker	Specific Marker	Data Set	PG Marker
Non-monetary markers	Number of different games played	Bwin	Account closure (9) RG flags (19)
	Average duration of a session	Win2day	Self-exclusion (17)
	Number of days an individual plays	Bwin	Self-exclusion (4)
	Total number of bets	Bwin	RG flags (20)
Monetary markers	Variability of bet size	Bwin	RG flags (19)
	Net loss	GTECH	Self-exclusion (13)
	Net loss	Paf (Finland)	Unprompted limit setting, limit removal (40)
	Increasing net loss over time	Bwin	Account closure (22)
	Increasing bet amount over time	Bwin	Account closure (22)