A guide to datasets used in Division on Addiction/bwin.party publications

This document generally describes the datasets that provided the evidence base for a series of publications focusing on actual Internet gambling activity. These datasets derive from the collaborative Internet gambling research project between the Division on Addiction (Division) and bwin.party digital entertainment (***bwin.party***), an Internet betting service provider headquartered in Vienna, Austria. This document and the referenced codebooks provide the information needed for accurate understanding and use of these datasets by researchers outside the Division. In addition to this description, we also have study specific descriptions for each data base within the available codebooks.

Below, we have organized our description of these datasets according to the primary cohorts of ***bwin.party*** subscribers whose betting behavior and other characteristics the datasets represent. In some cases, Division researchers have used a given cohort for a single publication. In other cases, Division researchers have used a given cohort for multiple publications. In each case, we make clear which cohort matches which publication(s).

Cohort 1. Eight months of betting behaviors for subscribers registered during February, 2005.

During 2006 the Division received a dataset that included eight months of real-time, Internet betting behaviors taking place from February 1, 2005 through September 30, 2005 for 43,851 subscribers who opened an account with ***bwin.party*** (previous names: betandwin.com, bwin.com) during the period from February 1, 2005 through February 28, 2005.

The following paper reports analyses of these data:

1. LaBrie, R. A., LaPlante, D. A., Nelson, S. E., Schumann, A., & Shaffer, H. J. (2007). Assessing the playing field: A prospective longitudinal study of Internet sports gambling behavior. *Journal of Gambling Studies*, 23, 347-362. In the following, we refer to this paper as "**LaBrie_2007"**.

We describe the specific inclusion/exclusion criteria for this publication in the codebook.

Cohort 2. Twenty four months of betting behaviors for subscribers registered during February, 2005.

During 2007, the Division received datasets with data that represented twenty-four months of aggregated betting behavior data for 48,114 subscribers who opened an account with ***bwin.party*** during the period from February 1, 2005 through February

28, 2005. We wrote the following papers based on our analyses of subscribers included in the 24-month cohort

1. LaBrie R. A., Kaplan, S. A., LaPlante, D. A., Nelson, S. E., and Shaffer, H. J. (2008). Inside the virtual casino: A prospective longitudinal study of actual Internet casino gambling. *European Journal of Public Health*, 18(4), 410-416. In the following, we refer to this paper as "**LaBrie_2008**".
2. LaPlante, D. A., Kleschinsky, J., LaBrie, R. A., Nelson, S. E., & Shaffer, H. J. (2009). Sitting at the Virtual Poker Table: A Prospective Epidemiological Study of Actual Internet Poker Gambling Behavior.*Computers in Human Behavior*, 25(3). 711-717. In the following, we refer to this paper as "**LaPlante_2009**".
3. Braverman, J., & Shaffer, H. J.   (2010). How Do Gamblers Start Gambling: Identifying Behavioural Markers for High-risk Internet Gambling. *European Journal of Public Health*, 1-6. DOI:10.1093/eurpub/ckp232. In the following, we refer to this paper as "**Braverman_2010**".

Figure 1 illustrates the interrelationships among the analytic datasets underlying these publications.
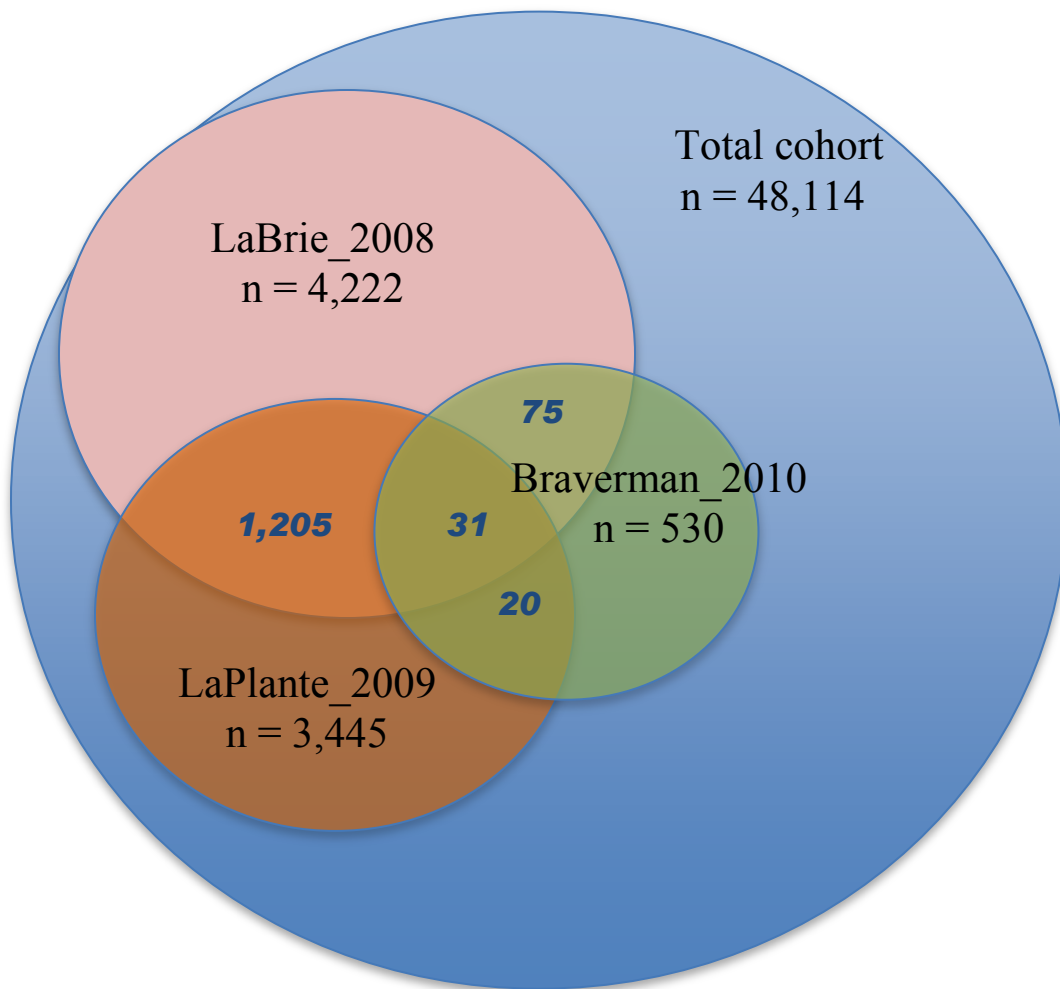
- The analytic dataset for LaBrie_2008 includes 4,222 subscribers who had casino betting records (i.e., they played casino-type games during the window of observation).
- The analytic dataset for LaPlante_2009 includes 3,445 subscribers who had poker betting records.
- The analytic dataset for Braverman_2010 includes 530 subscribers who had live action sports betting records for the first month since their deposit date and closed their account with **bwin.party**

There are varying degrees of overlap among these three analytic datasets.

- All three datasets share 31 subscribers.
- The analytic datasets for Braverman_2010 and LaBrie_2008 have 106 subscribers in common.
- The analytic datasets for Braverman_2010 and LaPlante_2009 have 51 subscribers in common.
- The analytic datasets for LaBrie_2008 and LaPlante_2009 have 1236 subscribers in common.

We describe the specific inclusion/exclusion criteria for each individual publication within each publication's codebooks.

Figure 1. Interrelationships among the analytical datasets derived from a raw dataset representing twenty-four months of aggregated betting behaviors

Total cohort
n = 48,114

LaBrie_2008
n = 4,222

75

Braverman_2010
n = 530

1,205

31

20

LaPlante_2009
n = 3,445

Cohort 3. Responsible Gambling Cases and Control Cases: Entire betting history

During 2010, the Division received datasets that represented the entire betting history of 4,134 subscribers. Approximately half of these subscribers (i.e., 2,068) triggered a Responsible Gambling intervention between November 2008 and November 2009 (i.e., RG cases). The other half (i.e., 2,066) made an initial **bwin** deposit on the same day as a case, but did not trigger a responsible gambling intervention between November 2008 and November 2009 (i.e. controls).

The Division wrote the following papers based on analyses of these data:

1. Gray, H. M., LaPlante, D. A., & Shaffer, H. J. (2012). Behavioral characteristics of Internet gamblers who trigger corporate responsible gambling interventions. *Psychology of Addictive Behaviors, 26* (3), 527–535. In the following we refer to this paper as"**Gray_2012."**
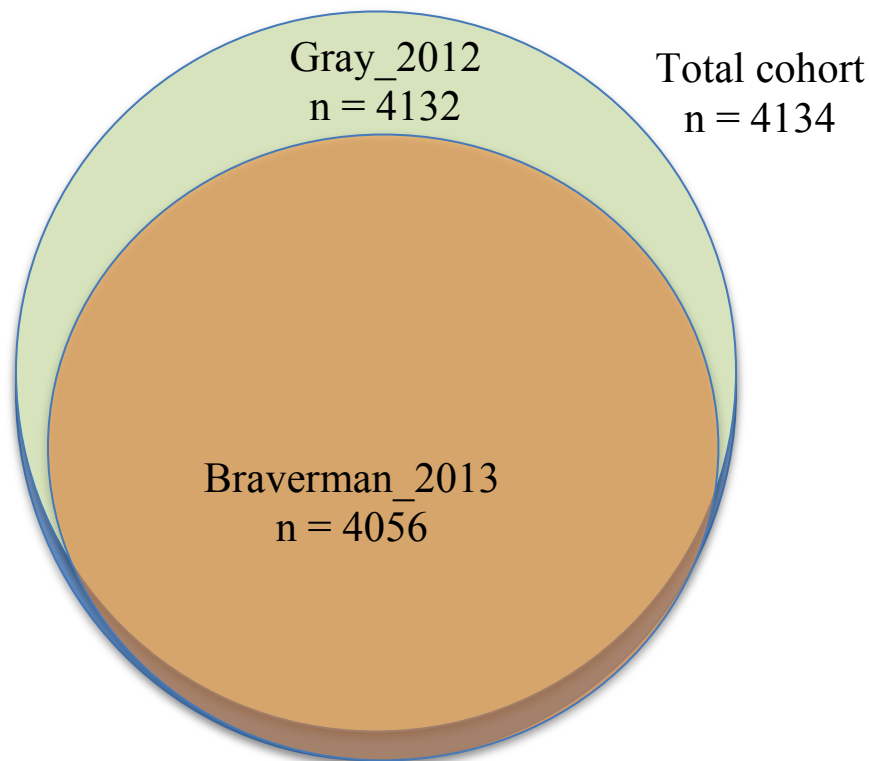
2.Braverman, J., LaPlante, D. A., Nelson, S. E., & Shaffer, H. J. (in press). Using Cross-game Behavioral Markers for Early Identification of High-risk Internet Gamblers *Psychology of Addictive Behaviors.* In the following we refer to this paper as "**Braverman_2013.**"

Figure 2 illustrates the interrelationships between the analytic datasets underlying these publications.

- Braverman_2013 describes 4,056 subscribers.
- These 4,056 subscribers are a subset of 4,132 subscribers described within Gray_2012 plus two participants who were excluded from Gray_2012.
- We describe the specific inclusion/exclusion criteria for each individual publication within each publication's codebooks.

We describe the specific inclusion/exclusion criteria for each individual publication within each publication's codebooks.

Figure 2. Interrelationships among the analytic datasets driven from a dataset representing entire betting history of RG and control cases

Interrelationships among cohorts

Figure 3 demonstrates the interrelationships among the three cohorts.

- Cohort 1 (Eight months of betting behaviors) is a subset of Cohort 2 (Twenty four months of betting behaviors) plus four subscribers who, for reasons unclear to us, were not included in **bwin.party's** Cohort 2 data delivery.
- Cohort 2 also includes 4267 subscribers who registered during February, 2005 but who are not included in Cohort 1. These subscribers did not engage in sports betting activity using non-promotional funds during the window of observation for Cohort 1 (i.e., February 1, 2005 through September 30, 2005). This is the likely reason **bwin.party** excluded this group from Cohort 1.
- Cohort 3: Responsible Gambling Cases and Controls Cases shares 62 subscribers with Cohort 1 and 66 with Cohort 2, all of whom are included in Cohorts 1 or 2 because they happened to register with **bwin.party** during February, 2005.
- 62 subscribers are shared by all three cohorts.

Figure 3 Interrelationships among Cohorts 1-3

Cohort 2
n = 48,114

Cohort 1
n = 43,851

62

4

Cohort 3
n = 4134