

Simple construct evaluation with latent class analysis: An investigation of Facebook addiction and the development of a short form of the Facebook Addiction Test (F-AT)

Michael Dantlgraber¹ · Eunike Wetzel¹ · Petra Schützenberger² · Stefan Stieger¹ ·
Ulf-Dietrich Reips¹

Abstract In psychological research, there is a growing interest in using latent class analysis (LCA) for the investigation of quantitative constructs. The aim of this study is to illustrate how LCA can be applied to gain insights on a construct and to select items during test development. We show the added benefits of LCA beyond factor-analytic methods, namely being able (1) to describe groups of participants that differ in their response patterns, (2) to determine appropriate cutoff values, (3) to evaluate items, and (4) to evaluate the relative importance of correlated factors. As an example, we investigated the construct of *Facebook addiction* using the Facebook Addiction Test (F-AT), an adapted version of the Internet Addiction Test (I-AT). Applying LCA facilitates the development of new tests and short forms of established tests. We present a short form of the F-AT based on the LCA results and validate the LCA approach and the short F-AT with several external criteria, such as chatting, reading newsfeeds, and posting status updates. Finally, we discuss the benefits of LCA for evaluating quantitative constructs in psychological research.

Keywords Latent class analysis · Bifactor model · Internet addiction · Facebook · Short form

✉ Michael Dantlgraber
michael.dantlgraber@uni-konstanz.de

¹ Department of Psychology, University of Konstanz, Box 31, 78457 Konstanz, Germany

² Department of Psychology, University of Vienna, Vienna, Austria

Most psychological constructs, including personality traits, cognitive abilities, interests, or attitudes, are assumed to be of a quantitative nature. That is, the construct can be thought of as a continuous dimension, with people possessing varying levels of the construct. In consequence, the main focus in psychological research and test development is on methods that capture the quantitative nature of the construct.

One example of a widely applied method is factor analysis, which explains the covariation among a larger number of observed variables by using a smaller number of latent variables (e.g., Kahn, 2006). Although the basic principle of factor analysis is the maximization of explained variance, it cannot differentiate between valid common variance among factors assessing the target construct and common variance that is unrelated to the target and that exists due to shared method variance. To separate the two sources of variance, the multitrait, multimethod approach was developed (Campbell & Fiske, 1959). For example, a paper–pencil version of a particular test can be compared with an online version of the same test and with an offline computerized version (for a discussion, see Reips, 2006). However, the estimated factors may still be biased if the methods used for data collection are imbalanced; that is, some methods are more similar than others. The multitrait, multimethod approach thus needs to be complemented with data analysis methods that allow further insights into the construct under investigation—that is, a *multi-analysis approach*.

The aim of this study is to illustrate how latent class analysis (LCA; Lazarsfeld, 1950; Lazarsfeld & Henry, 1968), a method originally developed to investigate qualitative differences between groups of respondents, can be applied to the investigation of quantitative constructs as one option in a multi-analysis approach. We demonstrate the possibilities of LCA and its benefits over factor-analytic methods using the recently developed construct *Facebook addiction*. In the following section, we describe in more detail the method of LCA

and its advantages. Then we describe the construct of Facebook addiction, which originated from the investigation of Internet addiction.

Latent class analysis

LCA (Lazarsfeld, 1950; Lazarsfeld & Henry, 1968) explains interindividual differences in response patterns by means of a given number of latent classes (subgroups of participants). LCA estimates the size of the classes and a membership probability for each participant within each class. Participants can be manifestly allocated to the latent class for which their class membership probability is highest. The estimated classes are disjunctive and exhaustive, and each one is defined by a specific pattern of category probabilities for each item. The exact number of latent classes is specified by the researcher. In most applications, LCAs with varying numbers of latent classes are compared to each other regarding their fits to the data. Which number of classes is appropriate for the description of the data can be judged using information criteria such as the Bayesian information criterion (BIC; Schwarz, 1978) and the consistent Akaike information criterion (CAIC; Bozdogan, 1987). LCA estimates the specific pattern of category probabilities without any a priori assumptions about the nature of the classes. The estimated solution can describe qualitative and also quantitative aspects of the data. For a more technical and detailed introduction to LCA, see McCutcheon (1987) and Hagenaaers and McCutcheon (2002).

LCA can be applied to the same data as exploratory factor analysis (EFA), but it is important to note that EFA can only describe the quantitative elements of the data. Therefore, LCA is even less restrictive and more exploratory than EFA. In previous research, LCA has been applied to identify different subgroups of youths that differ in their targeted communication about substance abuse (Kam, 2011) or different types of musical consumers (Chan & Goldthorpe, 2007).

LCA can also be used for confirmatory purposes. If LCA shows a structure that confirms the hypothesis made before, this indicates that conducting more restrictive analyses such as confirmatory factor analysis is appropriate.

In the past, LCA and factor analysis have been combined to create latent class factor analysis (LCFA; Kankaraš & Moors, 2009; Moors, 2004). LCFA tests whether groups of participants react in distinctively different ways to a given item pool. This method is used “for investigating measurement equivalence and detecting response bias” (Kankaraš & Moors, 2011). Mixed Rasch models can be used for the same purpose. For example, various response styles can be investigated (Rost, Carstensen, & von Davier, 1997; Wetzels, Carstensen, & Böhnke, 2013).

In the present article, we present an approach that works without needing to combine LCA and factor analysis into a single method. Instead, we use LCA as an internally created criterion for investigating multifacet constructs that are usually summed up to a single score. We use LCA to investigate specific sources of variance and their need to be represented in a measurement tool. We then test our approach by investigating the Facebook Addiction Test (F-AT), an adapted version of the Internet Addiction Test (I-AT; Young, 1998a, b), and by developing a short form of the F-AT.

Internet addiction and Facebook addiction

The term *addiction* is usually associated with substance abuse, but “interest in the behavioral addictions has grown in the past decade” (Black, Kuzma, & Shaw, 2012, p. 345). Although Grant, Brewer, and Potenza (2006) showed that substance addictions and behavioral addictions share many characteristics (see also Griffiths, 2005), Black et al. (p. 346) recommended viewing these disorders “as different and unique behavioral expressions of addictions.” The construct *Internet addiction* is one of these behavioral addictions. It describes the problematic use of the Internet and can be measured by the Internet Addiction Test (I-AT; Young, 1998a, b). The I-AT consists of 20 items that are rated on a 5-point Likert scale ranging from 1 (*not at all*) to 5 (*always*). The I-AT contains items such as *How often do you feel preoccupied with the Internet when offline*, or *How often do you fantasize about being online?*, or *How often does your job performance or productivity suffer because of the Internet?* Studies investigating the factor structure of the I-AT have found different numbers of latent factors, ranging from a general factor (Khazaal et al., 2008; Korkeila, Kaarlas, Jääskeläinen, Vahlberg, & Taiminen, 2010) up to six factors (Ferraro, Caci, D’Amico, & Di Blasi, 2007; Widyanto & McMurrans, 2004). Unfortunately, the results of these studies are difficult to compare, because they were conducted in different countries and used different versions of the I-AT.

To gain a better understanding of the I-AT, Watters, Keefer, Kloosterman, Summerfeldt, and Parker (2013) applied bifactorial EFA, as recommended by Reise, Moore, and Haviland (2010). Bifactorial measurement models fit the data with one general factor and with additional, uncorrelated factors to model specific manifestations of the construct. Using the bifactorial approach, Watters et al. confirmed a two-factor structure reported earlier by Korkeila et al. (2010). Moreover, a two-factor structure was found by Barke, Nyenhuis, and Kröner-Herwig (2012) using a German version of the I-AT.

For the present investigation, we adapted the I-AT to Facebook use (hence, F-AT) and replaced the word *Internet*

by the word *Facebook*. Cam and Isbulan (2012) earlier had also adapted the I-AT to Facebook (Facebook Addiction Scale; FAS) using a 6-point Likert scale. For a detailed overview of social networking addiction and related constructs, including Internet addiction and Facebook addiction, see Griffiths, Kuss, and Demetrovics (2014). Predictors specifically of Facebook addiction are discussed in Hong and Chiu (2014) and Hong, Huang, Lin, and Chiu (2014).

In the following sections, we show the comparability of the two constructs Internet addiction and Facebook addiction, and in the next step, the additions of LCA to factor-analytic methods.

Method

Procedure and sample

A total of 1,019 participants responded to an online questionnaire. They were recruited via mailing lists and social networks and by posting the link to the online questionnaire on several research websites following procedures recommended by Reips and Birnbaum (2011). Only the data from participants who stated having a Facebook account were analyzed further ($n = 841$). Of these, 70 (8 %) had more than one account; 520 were women (62 %), 312 were men (37 %), and nine did not state their sex (1 %). A total of 354 participants were in a relationship (42 %), 338 were single (40 %), 122 were married (15 %), 16 were divorced (2 %), and 11 (1 %) did not state their relationship status. The participants' median number of Facebook friends was 230, and the mean number of friends was 325. Participants' mean reported age was 27.5 years ($SD = 9.12$). With respect to the participants' country of origin, 57 % reported being Austrian, 21 % reported being German, 9 % reported being US citizens, and 13 % reported other nationalities. The participants' median number of educational years was 15.0, and the mean number of educational years was 14.9 ($SD = 3.50$). At the end of the questionnaire, respondents had the opportunity to participate in a raffle for winning a voucher.

Measures

We adapted the F-AT from the original I-AT (Young, 1998a, b) by modifying the wording to the context of investigating Facebook behavior. Thus, for all 20 items, the word *Internet* was replaced by the word *Facebook*. Participants gave their responses on 5-point rating scales, as in the I-AT, ranging from 1 (*not at all*) to 5 (*always*). An English version and a German version of the F-AT were used. The total scale had a mean score of 30.24 ($SD = 9.87$; Cronbach's $\alpha = .92$).

In addition, we asked participants about the frequency of their engagement in several Facebook activities, such as chatting, reading newsfeeds, and posting status updates. For the activity ratings, we also used a 5-point Likert scale, ranging from *never* to *very often*.

Statistical analyses: Factor analysis and LCA

A factor analysis using promax rotation and two fixed factors was estimated, to investigate in how far the two-factorial structure of the I-AT that has been found in other research (Barke et al., 2012; Korkeila et al., 2010) could be confirmed for the F-AT we used. In addition, LCAs with one to six classes were estimated using the R (R Development Core Team, 2014) package *poLCA* (Linzer & Lewis, 2011). To decide which number of classes was appropriate for our data, we used the BIC (Schwarz, 1978) and the CAIC (Bozdogan, 1987). The characteristics of the classes will be described using profiles of the expected means (Kempf, 2012), the sizes of the classes, and χ^2 tests of class membership with criteria such as sex and marital status. The expected means were calculated by summing the product of the category probabilities with the respective category values. On the basis of line profiles of the expected means and sizes of the classes, cutoff values of the quantitative scores can be derived, and items and correlated factors can be evaluated.

Results

In the following section, the results on the factor analysis of the F-AT data will be reported first. Subsequently, the results of the LCA will be reported. We will show (1) how to describe the characteristics of the classes, (2) how to determine cutoff values with LCA, (3) how to evaluate items with LCA, and (4) how to evaluate the importance of correlated factors. Finally, we will report (5) how F-AT scores are related to frequency ratings for different Facebook activities and (6) how LCA can be applied in the development of short instruments.

Using factor analysis and promax rotation, we found the same two-factor structure that Barke et al. (2012) had found for Internet addiction. Therefore, we used their factor labels by adapting them to the context of Facebook addiction. Only one noteworthy difference was found. In contrast to Barke et al.'s findings, the item *How often do you lose sleep due to late-night Facebook log-ins?* had a higher loading on Factor 1, *Emotional and cognitive preoccupation with Facebook*, than on Factor 2, *Loss of control and interference with daily life*. Nevertheless, this finding is probably caused by the different samples and not by a substantial difference between the constructs Internet addiction and Facebook addiction. The item allocation to the factors that we found here was later used for structural equation modeling (SEM; see the "Evaluation of the importance of correlated factors with LCA" section).

LCA results

For further evaluation of the construct Facebook addiction, we applied LCA. The lowest values for different information criteria (BIC, CAIC) indicate the best compromise between parsimony and data explanation. Table 1 shows that both BIC and CAIC suggested a division of the participants into three classes. For this reason, we only report analyses on this classification in the following sections.

Characteristics of the classes The interpretation of the three classes is based on (1) class sizes and (2) line profiles of the expected means of the items. The first class contained 44 %, the second class contained 39 %, and the third class contained 17 % of the participants. The average class membership probabilities for these three classes were .95 for Class 1, .95 for Class 2, and .97 for Class 3. Thus, respondents could be allocated to a particular class with a very high degree of certainty. Figure 1 shows the 20 items of the F-AT on the *x*-axis, ordered by average item difficulty. The *y*-axis shows the expected means for each group. The expected means can be interpreted as regular means; hence, Fig. 1 shows the means for each item and each class. In Fig. 1, each group is represented by a single line.

The line profiles of expected means for the three classes indicate a qualitatively distinct structure. In addition to that qualitative result, the line for Class 1 lies completely under the line for Class 2, and the line for Class 2 lies completely under the line for Class 3. That means that the estimated classes can be ordered quantitatively; that is, ordinal homogeneity exists (Kempf, 2012).

The first class (44 %, the lowermost line) contains participants with expected means near 1, indicating that they rated most items using 1 (*not at all*). The participants in this class therefore appear to use Facebook in a very moderate way. The second class (39 %, the middle line) contains participants that endorsed categories higher than a 1 on some items. The third class (17 %, the upper line) contains participants who rarely

endorsed a 1 (*not at all*). This class has expected means around the middle of the response scale for most items, and endorsed near the top of the scale on only a few items. Nevertheless, as compared to the other two classes, the respondents allocated to this class may tend to use Facebook in a way that could interfere with their daily life and responsibilities. Therefore, we call Class 3 the *risk* class.

The three classes consisted of approximately equal proportions of men and women [χ^2 test for sex and class membership: $\chi^2(2) = 2.73, p = .26$]. In contrast, there was a significant relationship between relationship status and class membership, $\chi^2(6) = 29.35, p < .001$, with a higher proportion of singles being allocated to the risk class (23 %) than were participants in a relationship (14 %) or married participants (15 %).

Determining cutoff values with LCA

The expected means of all three classes were low relative to the 5-point Likert scale that is used in the F-AT (see Fig. 1), and at a first glance participants in none of the three classes seem to use Facebook in a problematic way. This interpretation of the data would be correct if a rating in the middle of the response scale were equivalent to unproblematic Facebook behavior. However, it is necessary to consider that although the F-AT does not really measure a clinical addiction, the items resemble a checklist of negative symptoms. Therefore, participants that do not or that rarely show any problematic Facebook behaviors have ratings of 1 (*not at all*) on many items. As we mentioned above, such participants are represented in Classes 1 and 2. The participants in Class 3 have expected means above 1.5 on every item, indicating that they rarely rate 1 (*not at all*) on any of the items. For this reason, the Facebook use of the participants allocated to the latter class can be interpreted as potentially problematic.

The best way to identify participants that belong to the relevant Class 3 is to use LCA. LCA estimates three membership probabilities for each single participant, to each of the three classes. The participants can then be allocated to one of the classes according to their highest membership probability. Estimations were made using WINMIRA (von Davier, 2001), which resulted in 3.9 % false positives and 0.5 % false negatives in our sample. The quantitative ordering of the classes (see Fig. 1) reduced or eliminated their score overlap. Therefore, it was possible to estimate a cutoff criterion with LCA for the administration of the F-AT in practice, where there is no appropriate sample for estimating an LCA. Cutoff values can be estimated in the test validation process, and the classification of respondents is then possible simply by calculating the F-AT scores.

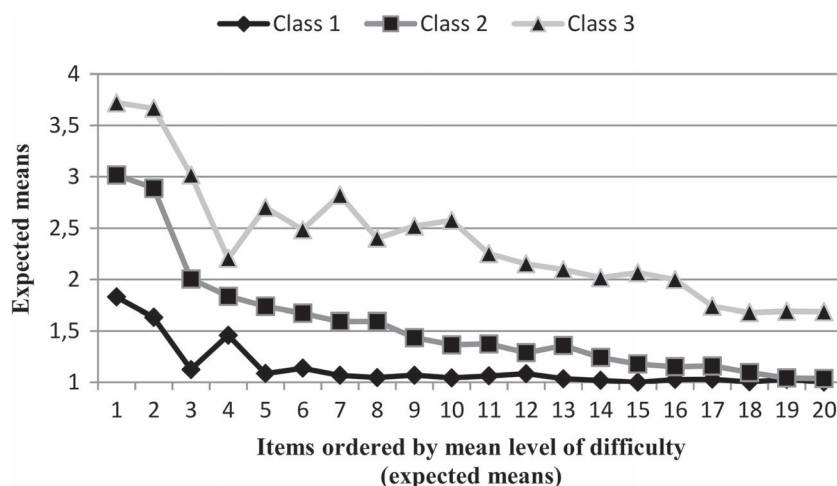
By considering the size of the relevant Class 3 (17 %), we calculated the 83rd percentile of the score distribution in our sample (100 %–17 %). This corresponds to a total score of 39. Participants who have scores above 39 can be

Table 1 Comparison of latent class models with different numbers of classes

Number of Classes	Log Likelihood	Number of Parameters	BIC	CAIC
1	-14,302.42	79	29,136.88	29,215.88
2	-12,340.68	159	25,752.15	25,911.15
3	-11,822.59	239	25,254.75	25,493.75
4	-11,598.45	319	25,345.24	25,664.24
5	-11,383.59	399	25,454.27	25,853.27
6	-11,289.45	479	25,804.76	26,283.76

BIC = Bayesian information criterion, CAIC = Consistent Akaike information criterion. The best-fitting model is depicted in boldface

Fig. 1 Expected means for each class and item: Class 1 (44 %), Class 2 (39 %), Class 3 (17 %)



allocated to the risk class. Regarding the membership probabilities of the LCA, this procedure would result in 4.7 % false positives and 2.0 % false negatives in our sample. The procedure is similar to the approach used by Demetrovics et al. (2012) and Pápay et al. (2013).

Item evaluation with LCA

LCA is very useful for evaluating items that were constructed for measuring a quantitative construct. We will demonstrate this by using Items 4 and 7 in an example (see Fig. 1). The fourth item, *How often do you form new relationships with Facebook users?*, does not differentiate well between the classes, as the expected means of all classes are close together. In contrast, Item 7, *How often do you check Facebook before something else that you need to do?*, differentiates well because the expected means differ strongly. The obvious content difference between the two described items is also expressed in the analysis shown in Fig. 1.

Usually, item–test correlations or factor loadings are used for item comparisons, to discern their ability to differentiate. The results are comparable but not identical to the results of the LCA (see the next section). But item–test correlations and factor loadings do not depend only on item quality, but also on item difficulty. Items that have a medium level of difficulty show the highest item–test correlations. However, a test that is well constructed consists of items with low and high difficulties, in order to adequately represent a construct’s whole continuum. Therefore, it is important to evaluate an item regarding both aspects, namely its ability to differentiate and its level of difficulty. Using LCA, both aspects can be depicted in a single figure. Furthermore, conclusions based on item–test correlations or factor loadings depend to a greater degree on the investigated and potentially biased item pool than do conclusions based on LCA results. We elaborate on this aspect in the next section.

Evaluation of the importance of correlated factors with LCA

LCA can be used for ranking correlated factors in terms of their importance. This is especially useful if the goal is to investigate a relatively new construct such as Facebook addiction, for which little previous research exists.

Constructs are often modeled by correlated factors in order to give consideration to their different aspects. The fact that the estimated factors have something in common with the other factors complicates the interpretation of the factor loadings. It is not clear to which extent a high factor loading is sourced by the common part or the specific part of the estimated factor. That is one of the reasons why Reise et al. (2010) argued for bifactor modeling, a method that differentiates these two sources of variance. In the following discussion, we compare three different confirmatory factor models of the F-AT using SEM and discuss the additions of LCA.

Table 2 lists the items of the F-AT in the same order as in Fig. 1 and compares the factor loadings of a simple one-factor model (i.e., one general factor is extracted) with the factor loadings of the correlated factor model and the bifactor model (i.e., one general factor and two additional, uncorrelated factors are estimated).

The general-factor model estimates some high loadings for both groups of items. The loadings of the first group of items (those connected to Factor 1) have a mean loading of .60 and range from .33 to .70. The loadings of the second group of items, which are connected to Factor 2, have a mean loading of .66 and range from .58 to .74.

The correlated factor model tends to result in higher loadings because it takes the specific sources of variance into consideration. The problems of the correlated factor model can be shown in comparison to the bifactor model. For example, Items 5 and 10 have the same loadings in the correlated

Table 2 Factor loadings of the Facebook addiction test items depicted for different factor models

Item Number in Fig. 1	Factor Loadings					
	General Factor	Correlated Factors		Bifactor Model		
		F1	F2	Core Factor	F1	F2
1	.62	–	.71	.66	–	.19
2	.58	–	.68	.61	–	.29
3	.72	–	.79	.75	–	.19
4	.33	.34	–	.30	.15	–
5	.64	–	.74	.62	–	.62
6	.65	.64	–	.63	.20	–
7	.74	–	.72	.78	–	–.06
8	.62	–	.73	.62	–	.50
9	.70	.69	–	.67	.23	–
10	.73	–	.74	.76	–	.06
11	.60	.63	–	.53	.31	–
12	.63	.63	–	.56	.27	–
13	.59	.61	–	.51	.33	–
14	.67	.72	–	.57	.43	–
15	.70	.69	–	.64	.27	–
16	.62	.67	–	.53	.39	–
17	.50	.53	–	.41	.33	–
18	.58	.66	–	.43	.58	–
19	.58	.64	–	.46	.47	–
20	.63	.71	–	.48	.61	–

F1 = Factor 1, *Emotional and cognitive preoccupation with Facebook*; F2 = Factor 2, *Loss of control and interference with daily life*

factor model (.74), but the loading of the former is sourced by both the common-variance and specific-variance parts, and the loading of the latter is only sourced by the common-variance part. In general, the bifactor model provides more detailed information than the correlated factor model.

Here we call the common factor of the bifactor model the *core factor*, because it is not the same as the general factor of a simple one-factor model, but is corrected for the specific variance sources. The core-factor loadings of the first group—items that are connected to Factor 1—are lower than the general-factor loadings, since they have a mean loading of .52 and range from .30 to .67. The core-factor loadings of the second group of items are higher than the general-factor loadings, since they have a mean loading of .69 and range from .61 to .78. That means that the Factor 2 items represent the core of the construct better than the Factor 1 items do. Furthermore, it shows that the general factor is biased in the direction of Factor 1, because of the imbalanced item pool that consists of 13 Factor 1 items and only seven Factor 2 items. It therefore overrepresents the specific source of variance of Factor 1.

It is obvious that the bifactor model surpasses the other models. But the bifactorial approach cannot answer the questions of whether the specific sources of variance are useful to identify people that belong to the risk group or whether it is useful to have items in the pool with substantial specific loadings. Comparisons to the (biased) general-factor model do not help. In the general-factor model, the principle of the maximization of explained variance makes every source of variance appear to be important if it is represented by enough items (the same applies for item–scale correlations).

For testing the use of the specific sources of variance, we used LCA as an internally created criterion that is not based on the principle of maximizing the explained variance and that combines the various sources of variance as the scale score does. Table 3 shows the correlations between the specific-factor loadings and the group differences of the expected means in the LCA while controlling for the core-factor loadings.

We found no indication in the data that the specific source of variance of the factor *Emotional cognitive preoccupation with Facebook* could be a useful indicator to differentiate between the risk class (Class 3) and Class 2, containing participants that are likely not concerned by Facebook addiction. But we did find an indication that the specific source of variance of the factor *Loss of control and interference with daily life* could help identify the risk class. The partial correlation between the loadings of the bifactor model and the differences between Classes 3 and 2 is substantial ($r = .48$). Both specific sources of variance hinder the differentiation between Class 2 and Class 1.

Only the highest correlation (–.78) reaches significance, but this is probably caused by the small number of items. Statistical tests view the items as a small sample of an underlying item pool, but of course the items represent the whole population of the test developed by Young (1998a, b). Conclusions about the specific loadings are based on a big sample of participants and can be drawn for the well-known I-AT and its adaptation (F-AT).

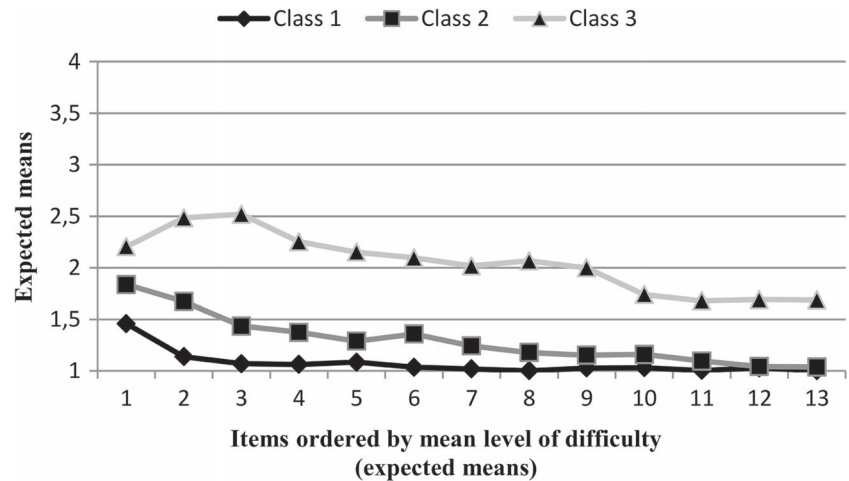
A graphical analysis also shows that a combination of the core and Factor 2 specific loadings is more useful for group

Table 3 Partial correlations between the expected mean differences of the latent classes and the specific loadings of the bifactor model, controlled for the core factor loadings

Differences of Expected Means	Specific Loadings of the Bifactor Model	
	Factor 1	Factor 2
EM C3 – EM C2	–.07	.48
EM C2 – EM C1	–.78*	–.61

EM = expected means; C = class; Factor 1, *Emotional and cognitive preoccupation with Facebook*; Factor 2, *Loss of control and interference with daily life*. * $p < .05$

Fig. 2 Factor 1 item comparison using LCA: Class 1 (44 %), Class 2 (39 %), Class 3 (17 %)



differentiation than is the combination of the core and Factor 1 specific loadings. Figures 2 and 3 show the lines profiles of the items separately for the two estimated factors. The line profiles represent the expected means of a single class, but of course there is variance within the classes. That means the overlap between the classes decreases when the line profiles diverge. Figure 2 shows the 13 items that have higher loadings on Factor 1, *Emotional and cognitive preoccupation with Facebook*, and Fig. 3 shows the seven remaining items that have higher loadings on Factor 2, *Loss of control and interference with daily life*.

A comparison of the two figures indicates that the seven items that comprise Factor 2 (Fig. 3) differentiate better between the quantitative groups than do the 13 items that make up Factor 1 (Fig. 2), because the line profiles of the expected means are farther apart for Factor 2. The line profiles of the expected means for Factor 1 are much closer together, and Classes 1 and 2 especially have practically identical (low) expected means across the 13 items. The differences between

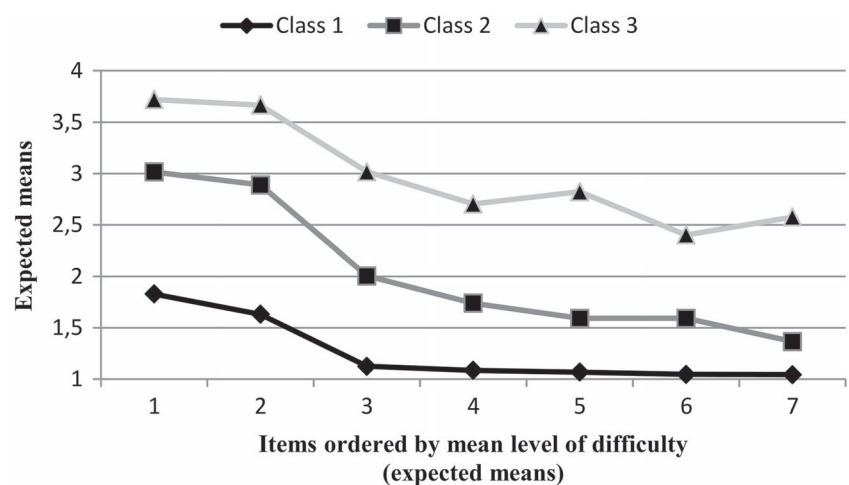
the classes are significantly larger for the Factor 2 than for the Factor 1 items. We estimated a Cohen's $d = 1.16$ ($p = .045$) for the comparison of Classes 3 and 2, and a Cohen's $d = 2.4$ ($p = .006$) for the comparison of Classes 2 and 1. Significance was tested using two-sided t tests.

We concluded that the second factor is the better indicator for the construct of Facebook addiction and is the more important factor, although the test consists of fewer Factor 2 items than Factor 1 items. Furthermore, we conclude that Factor 1 is not needed in the test. The core factor is better represented by the Factor 2 items, and there is no indication in the data that the specific source of variance of Factor 1 could be useful for a better group differentiation.

F-AT scores and Facebook activities

We tested our conclusions based on the internally created criterion (i.e., the LCA results) with an external criterion

Fig. 3 Factor 2 item comparison using LCA: Class 1 (44 %), Class 2 (39 %), Class 3 (17 %)



of Facebook addiction. For this purpose, we analyzed the relationship of F-AT scores to frequency ratings of engagement in several Facebook activities, such as chatting, reading newsfeeds, or posting status updates. Even if it is unclear which activities are especially connected to Facebook addiction, it is an obvious hypothesis to expect high F-AT raters to use Facebook more often in general. Table 4 shows the correlations between the total score (Cronbach $\alpha = .92$) and the activities, as well as the correlations between the two factor scales and the activities. The two scales have nearly the same reliability and consist of all items that are connected to the factors—that is, Scale 1 (Cronbach $\alpha = .88$) consists of 13 items (Fig. 2), and Scale 2 (Cronbach $\alpha = .88$) consists of seven items (Fig. 3).

All correlations in Table 4 are significant at $p < .05$. The correlations of the activities with Scale 2, which is based on seven items, are comparable to the correlations of the activities with the total scale, which is based on 20

Table 4 Correlations between Facebook activities and Facebook addiction scores

Facebook Activities	Total Scale	Scale 1	Scale 2
Reading news feeds	.21	.17	.22
Reading private messages	.16	.11	.18
Writing private messages	.22	.18	.23
Chatting with Facebook friends	.24	.21	.23
Posting on someone else’s Timeline	.33	.30	.31
Commenting on posts	.37	.33	.35
Reading posts	.30	.25	.30
Inviting people to own events	.21	.18	.22
Looking at content posted to a group	.27	.22	.28
Posting to a group	.25	.22	.25
Posting photos	.34	.32	.30
Looking at photos	.34	.28	.35
Commenting on photos	.40	.35	.39
Posting videos	.26	.24	.25
Looking at videos	.27	.24	.25
Commenting on videos	.31	.30	.28
Posting status updates	.35	.32	.32
Commenting on status updates	.37	.31	.38
Playing games on Facebook	.18	.25	.08
Updating basic profile information	.31	.32	.24
Using the Like button	.39	.32	.40
Sharing interesting content	.37	.34	.34
Looking at other profiles	.33	.28	.34

Scale 1 = Scale of the 13 Factor 1 items. Scale 2 = Scale of the seven Factor 2 items. All correlations are significant at the $p < .05$ level. Loadings are depicted in boldface when comparisons of the factor scale correlations are significant at $p < .05$ using Fisher’s Z transformation

items. Furthermore, the correlations between the activities and Scale 2 are descriptively higher than the correlations between the activities and Scale 1. The activities with a larger correlation for Scale 2 than for Scale 1 correspond to a ratio of 17:6. Under the null hypothesis of a ratio of 11.5:11.5, this is significant at $p < .05$. If the correlations are separately compared using Fisher’s Z transformation, the activity *Using the Like button* correlates higher with Scale 2 than with Scale 1, and the activity *Playing games*, which is loosely connected to the other Facebook activities, correlates higher with Scale 1 at $p < .05$. These results are further evidence that Factor 2, *Loss of control and interference with daily life*, is the better indicator and the more important component of Facebook addiction.

Generating short forms of tests with LCA

In the section above we showed how to determine the importance of correlated factors. We argued that Factor 2 is the better indicator and the more important factor. We showed that a scale consisting of all seven items loading on Factor 2 is comparable to the total scale (20 items) regarding its reliability and its correlations to external criteria. For this reason, we recommend the use of the Factor 2 scale as a short form of the F-AT (see Table 5).

The factor loadings of a one-factor model have a mean loading of .73 and range from .68 to .79. Using the same method described in the Determining Cutoff Values With LCA section, for the short form of the F-AT we calculated that participants who have total scores above 18 can be allocated to the risk class. Using the short form and the LCA classification results in 6.6 % false positives and 1.4 % false negatives, whereas using the short form and the cutoff criterion results in 4.0 % false positives and 3.2 % false negatives.

Table 5 Short form of the Facebook Addiction Test

Item Number in Fig. 1	Item	Factor Loading
1	How often do you find that you stay on Facebook longer than you intended?	.72
2	How often do you check Facebook before something else that you need to do?	.69
3	How often do you neglect household chores to spend more time on Facebook?	.79
5	How often does your job performance or productivity suffer because of Facebook?	.76
7	How often do you find yourself saying “just a few more min” when on Facebook?	.68
8	How often does your work suffer because of the amount of time you spend on Facebook?	.75
10	How often do you try to cut down the amount of time you spend on Facebook?	.71

Discussion

In this article, we reported on the development and analysis of the Facebook Addiction Test (long and short forms) using LCA. We showed how valuable information on the construct and its assessment can be obtained from LCA to complement the results of factor analysis—namely, support for the decision about the sources of variance that should be represented in the test.

In a first step, we replicated the structure of two correlated factors found by Barke et al. (2012) for the I-AT. For this reason, we argue that the Facebook addiction construct is comparable to the Internet addiction construct, only that the former is narrower, applying to the specific context of Facebook, whereas the latter is broader, applying to the Internet in general. The comparability of the constructs does not imply that the results of the F-AT and the I-AT are equivalent. Because Facebook is a part of the Internet, participants who indicate a problematic use of Facebook would indicate a problematic use of the Internet as well, but a problematic use of the Internet does not necessarily indicate a problematic use of Facebook. Future research could adapt the I-AT to other services and social media besides Facebook and investigate to what extent similar underlying factors exist and whether similarly sized groups of other social media users face light to moderate risk of addiction.

We showed that a division of Facebook users into three classes using LCA is reasonable and empirically justified. The low expected means of the classes imply that the participants of the third class (17%) are the only ones who tended to use Facebook in a problematic manner that might interfere with their daily lives. Being able to quickly identify and focus on people who fall into this risk class from test results will be useful for further research, as well as for clinical applications of the F-AT. Thus, the F-AT as presented here, and in particular its short form (see below), could be used as a screening instrument.

Quantitative scores on behavioral questionnaires are often used for classifying participants by determining cutoff values, especially in clinical assessment. Naturally, classification is a qualitative process. We showed how LCA, as a qualitative method, can be used to support the classification decision when determining appropriate cutoff values of quantitative scores. Cutoff values can be estimated in the test validation process and then can be used for applications of the test in single cases.

In line with Demetrovics et al. (2012) and Pápay et al. (2013), we compared the accuracy of classification using cutoff values with the accuracy of classification using LCA. This can be used for detecting qualitative constructs for which quantitative cutoff estimation is inappropriate. Determining cutoff values is reasonable if the classes recommended by the information criteria show a quantitative structure. That is the case when the line profiles of the expected means representing the classes do not overlap (i.e., when ordinal homogeneity exists; Kempf, 2012).

Overlapping lines indicate qualitative aspects of the investigated item pool. Content considerations then become necessary, and result in two options. If the particular items are *essential* with regard to content coverage, overlapping lines indicate a qualitative construct that is not adequately represented by a quantitative score. Researchers then should use qualitative variables instead. If the particular items are *not essential* with regard to content coverage, overlapping lines indicate suboptimal items of a quantitative construct. Hence, LCA can be used to identify suboptimal items.

LCA can also be used for comparing items that are *not* responsible for overlapping lines—that is, items that differentiate well. We showed that some items differentiate considerably better between the groups than do other items. In this regard, LCA can be used for comparing specific sources of variance that can be modeled in a bifactor model. Bifactorial modeling is hampered by the problem that it does not deliver any indication whether the specific loadings are useful and should be represented in the test. For this reason, the results of an LCA are valuable additions to the results of a factor analysis.

We used LCA as an internally created criterion that is based on few assumptions (e.g., it treats the categories as a nominal scale during estimation and does not require normally distributed data) and that combines the various sources of variance as the scale score does. The LCA approach is not based on maximization of explained variance and surpasses item–scale correlations or the loadings of the general model because it is less influenced by an imbalanced item pool.

We showed that the items loading on Factor 2, *Loss of control and interference with daily life*, have higher core-factor loadings in the bifactor model, and that there are no indications in the data that the specific loadings of Factor 1, *Emotional and cognitive preoccupation with Facebook*, are useful for better group differentiation. For this reason, we argue that Factor 2 is the better indicator and the more important component of the construct Facebook addiction.

In addition, we showed that a single scale only consisting of the seven Factor 2 items shows validity similar to that of the complete 20-item scale. Because the items of Factor 1 do not differentiate as well between the classes and do not increase external validity, we argue that the test can be reduced to Factor 2, thus ignoring Factor 1. Following this argumentation, the Factor 2 scale can be used as a short form of the F-AT.

One may question the development of a short test version by simply ignoring a facet of a construct. Often the goal is to represent all facets of a construct when developing a short form of a published test. For example, Pápay et al. (2013) presented a short form of the Problematic Online Gaming Questionnaire (POGQ-SF) that consists of 12 items representing six facets. We

believe representing all facets of a test for problematic behavior in a short form is only necessary if all facets (i.e., all specific loadings) are useful for identifying the relevant group of participants. Using latent profile analysis of correlated factors facilitates factor comparisons but cannot answer questions regarding the need for the facets. To arrive at answers, one has to differentiate the common-variance part and the specific-variance part.

However, we think it is important to clearly separate the construct and the application of the test. Ignoring a construct facet in the needed application of a derived test does not mean a redefinition of the construct. A certain behavioral pattern (e.g., Facebook use) could be associated with a construct (e.g., Facebook addiction) but not be suitable as an indicator in a questionnaire, because it is also associated with other constructs. The presented multi-analysis approach supports Griffiths' (2010) assumption from a study of two cases that excessive behaviors should be distinguished from addiction.

As we mentioned above, several studies have investigated the factor structure of the I-AT, but more research will be needed to investigate the constructs Internet addiction and Facebook addiction, as well as related constructs (see Griffiths et al., 2014). Using LCA in this study was only a further step toward arriving at a better understanding of behavioral addictions related to Internet use.

Limitations

The approach for determining cutoff values presented here can be used for examining existing cutoff values—for example, in clinical research. Note, however, that the cutoff values we computed for our sample may not generalize to other—in particular, to clinical—samples. Thus, future research could apply LCA to F-AT data from a clinical sample and test whether the cutoff values reported here hold for this sample.

Furthermore, our conclusions about the use of the specific sources of variance can be drawn for the investigated item pool (i.e., for the F-AT) and not necessarily for the whole construct of Facebook addiction.

As we mentioned above, using several Facebook activities is not an optimal external criterion for Facebook addiction. On the other hand, it is often impossible or very costly to collect data that can undoubtedly be used as valid external criteria. Thus, securing conclusions based on analyzed external criteria by comparing them with internal LCA results is reasonable.

In this article, we showed what an additional LCA can add to the interpretation of a frequently used method, namely factor analysis. We used LCA as an internally created criterion for investigating multifaceted constructs that are usually summed up to a single score. That means we were

investigating and compressing scales that are not unidimensional (Rasch, 1960, 1961). For future research, it would be interesting to compare compressed scales of multifaceted constructs that focus on all factors that are useful for diagnostic purposes, with very restrictive unidimensional scales regarding several external criteria. It is difficult to assess whether the advantages of unidimensional scales can compensate for the lack of flexibility during item selection. The demands of the analysis method presented in this article are more restrictive than the standard demands of classical test theory, but still they are less restrictive than the demands of probabilistic test theory (see Kempf, 2012). For this reason, we think our analysis is a helpful addition to the toolbox available to researchers and test developers.

References

- Barke, A., Nyenhuis, N., & Kröner-Herwig, B. (2012). The German version of the internet addiction test: A validation study. *CyberPsychology, Behavior, and Social Networking, 15*, 534–542. doi:10.1089/cyber.2011.0616
- Black, D. W., Kuzma, J., & Shaw, M. (2012). Unique consequences of behavioral expressions of addiction. In H. J. Shaffer, D. A. LaPlante, & S. E. Nelson (Eds.), *APA addiction syndrome handbook, Vol. 1: Foundations, influences, and expressions of addiction* (pp. 329–351). Washington, DC: American Psychological Association.
- Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika, 52*, 345–370. doi:10.1007/BF02294361
- Cam, E., & Isbulan, O. (2012). A new addiction for teacher candidates: Social networks. *Turkish Online Journal of Educational Technology, 11*, 14–19.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105. doi:10.1037/h0046016
- Chan, T. W., & Goldthorpe, J. H. (2007). Social stratification and cultural consumption: Music in England. *European Sociological Review, 23*, 1–19. doi:10.1093/esr/jcl016
- Demetrovics, Z., Urbán, R., Nagygyörgy, K., Farkas, J., Griffiths, M. D., Pápay, O., . . . Oláh, A., (2012). The development of the Problematic Online Gaming Questionnaire (POGQ). *PLoS ONE, 7*, e36417. doi:10.1371/journal.pone.0036417
- Ferraro, G., Caci, B., D'Amico, A., & Di Blasi, M. (2007). Internet addiction disorder: An Italian study. *CyberPsychology and Behavior, 10*, 170–175. doi:10.1089/cpb.2006.9972
- Grant, J. E., Brewer, J. A., & Potenza, M. N. (2006). The neurobiology of substance and behavioral addictions. *CNS Spectrums, 11*, 924–930. doi:10.1017/s109285290001511x
- Griffiths, M. D. (2005). A 'components' model of addiction within a biopsychosocial framework. *Journal of Substance Use, 10*, 191–197. doi:10.1080/14659890500114359
- Griffiths, M. D. (2010). The role of context in online gaming excess and addiction: Some case study evidence. *International Journal of Mental Health and Addiction, 8*, 119–125. doi:10.1007/s11469-009-9229-x
- Griffiths, M. D., Kuss, D. J., & Demetrovics, Z. (2014). Social networking addiction: An overview of preliminary findings. In K. P. Rosenberg & L. C. Feder (Eds.), *Behavioral addictions. Criteria, evidence, and treatment* (pp. 119–141). New York, NY: Elsevier.

- Hagenaars, J. A., & McCutcheon, A. L. (2002). *Applied latent class analysis*. Cambridge, UK: Cambridge University Press.
- Hong, F.-Y., & Chiu, S.-L. (2014). Factors influencing Facebook usage and Facebook addictive tendency in university students: The role of online psychological privacy and Facebook usage motivation. *Stress and Health*. Advance online publication. doi:10.1002/smi.2585
- Hong, F.-Y., Huang, D.-H., Lin, H.-Y., & Chiu, S.-L. (2014). Analysis of the psychological traits, Facebook usage, and Facebook addiction model of Taiwanese university students. *Telematics and Informatics*, 31, 597–606. doi:10.1016/j.tele.2014.01.001
- Kahn, J. H. (2006). Factor analysis in counseling psychology research, training, and practice: Principles, advances, and applications. *Counseling Psychologist*, 34, 684–718. doi:10.1177/0011000006286347
- Kam, J. A. (2011). Identifying changes in youth's subgroup membership over time based on their targeted communication about substance use with parents and friends. *Human Communication Research*, 37, 324–349. doi:10.1111/j.1468-2958.2011.01408.x
- Kankaraš, M., & Moors, G. (2009). Measurement equivalence in solidarity attitudes in Europe. Insights from a multiple group latent class factor approach. *International Sociology*, 24, 557–579. doi:10.1177/0268580909334502
- Kankaraš, M., & Moors, G. (2011). Measurement equivalence and extreme response bias in the comparison of attitudes across Europe: A multigroup latent-class factor approach. *Methodology*, 7, 68–80. doi:10.1027/1614-2241/a000024
- Kempf, W. (2012). A pragmatic approach to Rasch-modeling: The loss of information index. *Institutional repository of the University of Konstanz*. Retrieved from <http://nbn-resolving.de/urn:nbn:de:bsz:352-209584>
- Khazaal, Y., Billieux, J., Thorens, G., Khan, R., Louati, Y., Scarlatti, E., & Zullino, D. (2008). French validation of the internet addiction test. *CyberPsychology and Behavior*, 11, 703–706. doi:10.1089/cpb.2007.0249
- Korkeila, J., Kaarlas, S., Jääskeläinen, M., Vahlberg, T., & Taiminen, T. (2010). Attached to the web—Harmful use of the internet and its correlates. *European Psychiatry*, 25, 236–241. doi:10.1016/j.eurpsy.2009.02.008
- Lazarsfeld, P. F. (1950). The logical and mathematical foundation of latent structure analysis. In S. A. Stouffer, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star, & J. A. Clausen (Eds.), *Studies in social psychology in world war II* (Measurement and prediction, Vol. 4, pp. 362–472). Princeton, NJ: Princeton University Press.
- Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Boston, MA: Houghton Mifflin.
- Linzer, D. A., & Lewis, J. B. (2011). PoLCA: An R package for polytomous variable latent class analysis. *Journal of Statistical Software*, 42, 1–29. Retrieved from www.jstatsoft.org/v42/i10
- McCutcheon, A. L. (1987). *Latent class analysis*. Newbury Park, CA: Sage.
- Moors, G. (2004). Facts and artefacts in the comparison of attitudes among ethnic minorities. A multi-group latent class structure model with adjustment for response style behaviour. *European Sociological Review*, 20, 303–320. doi:10.1093/esr/jch026
- Pápay, O., Urbán, R., Griffiths, M. D., Naggyörgy, K., Farkas, J., Kökönyei, G., . . . Demetrovics, Z. (2013). Psychometric properties of the problematic Online Gaming Questionnaire short-form and prevalence of problematic online gaming in a national sample of adolescents. *CyberPsychology, Behavior, and Social Networking*, 16, 340–348. doi:10.1089/cyber.2012.0484
- R Development Core Team. (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.R-project.org
- Rasch, G. (1960). *Studies in mathematical psychology: I. probabilistic models for some intelligence and attainment tests*. Oxford, UK: Nielsen & Lydiche.
- Rasch, G. (1961). *On general laws and the meaning of measurement in psychology*. Berkeley: University of California Press.
- Reips, U.-D. (2006). Web-based methods. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 73–85). Washington, DC: American Psychological Association.
- Reips, U.-D., & Birnbaum, M. H. (2011). Behavioral research and data collection via the Internet. In K.-P. L. Vu & R. W. Proctor (Eds.), *The handbook of human factors in Web design* (2nd ed., pp. 563–585). Mahwah, NJ: Erlbaum.
- Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, 92, 544–559. doi:10.1080/00223891.2010.496477
- Rost, J., Carstensen, C. H., & von Davier, M. (1997). Applying the mixed Rasch model to personality questionnaires. In J. Rost & R. Langeheine (Eds.), *Applications of latent trait and latent class models in the social sciences* (pp. 324–332). Münster, Germany: Waxmann.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464. doi:10.1214/aos/1176344136
- Von Davier, M. (2001). *WINMIRA 2001 [computer software]*. Kiel, Germany: Institute for Science Education.
- Watters, C. A., Keefer, K. V., Kloosterman, P. H., Summerfeldt, L. J., & Parker, J. D. A. (2013). Examining the structure of the internet addiction test in adolescents: A bifactor approach. *Computers in Human Behavior*, 29, 2294–2302. doi:10.1016/j.chb.2013.05.020
- Wetzel, E., Carstensen, C. H., & Böhnke, J. R. (2013). Consistency of extreme response style and non-extreme response style across traits. *Journal of Research in Personality*, 47, 178–189. doi:10.1016/j.jrp.2012.10.010
- Widyanto, L., & McMurrin, M. (2004). The psychometric properties of the internet addiction test. *CyberPsychology and Behavior*, 7, 443–450. doi:10.1089/cpb.2004.7.443
- Young, K. (1998a). *Caught in the net: How to recognize the signs of internet addiction and a winning strategy for recovery*. New York, NY: Wiley.
- Young, K. (1998b). Internet addiction: The emergence of a new clinical disorder. *CyberPsychology and Behavior*, 1, 237–244. doi:10.1089/cpb.1998.1.237