Binge-Watching, Self-Determination, and Well-Being

A Partially Successful Direct Replication and Extension of Granow et al. (2018)

Elena Erdmann[1] and Tobias Dienlin[2]

[1]Department of Media Psychology, University of Hohenheim, Germany

[2]Department of Communication, University of Vienna, Austria

Abstract

Binge-watching is a highly relevant new media phenomenon. An increasing number of people watch multiple episodes of their favorite series online, a process often spanning several hours. Granow et al. (2018) showed that binge-watching is accompanied by both positive as well as negative psychological experiences. Because it is one of the most important studies on the topic, we conducted a direct replication of Granow et al. (2018). We also extended the original study by including additional variables derived from self-determination theory. We conducted an online questionnaire with 668 respondents. Results showed that we could not reproduce the exact model that was reported. We also needed to adapt several of the originally tested measures to achieve satisfactory model fit. After introducing these changes, results showed that we replicated 67% of all effects in terms of significance (i.e., originally significant effects were also significant in the replication, and vice versa). When looking at effect size (i.e., originally reported effects fell into the replication's 95% confidence interval), we replicated 42%. Our study represents a partial replication of Granow et al. (2018). In our extension, we found two further positive links between binge-watching and well-being: Binge-watching was associated not only with greater autonomy, but also with increased feelings of competence and relatedness. Competence and relatedness were related to increased well-being, which included vitality, recovery experience, and media enjoyment. In line with Granow et al. (2018), we found that binge-watching is related to self-determined media use and improved psychological well-being.

Keywords: binge-watching, streaming, well-being, self-determination theory, replication, structural equation modeling

Video-on-demand services such as Netflix or Amazon Prime Video are increasingly replacing linear television – they have become a central component of digital entertain- ment (Koch & Beisch, 2020). Due to the constant availability and almost infinite variety of series and movies, a new form of media entertainment called "binge-watching" has developed. Binge-watching differs substantially from traditional television. Compared to linear television, binge-watching is longer, more intense, and more self-determined (Granow et al., 2018, p. 393). Following Granow et al. (2018, p. 393), binge-watching is understood as the intense and consecutive consumption of series in a single sitting.

Although media entertainment in general is often linked to improved well-being, binge-watching in particular is associated primarily with risks such as addiction and impaired mental health (Birch, 2019). Granow et al. (2018), however, were one of the first to paint a more nuanced picture, demonstrating that intensive series consumption can be both negative and positive for a user's well-being.

The study by Granow et al. (2018) is highly cited, and it rightly belongs to one of the most important and seminal studies on binge-watching. In general, to establish that scholarly findings are reliable and robust, they should routinely be scrutinized and replicated (Merton, 1974) – even more so if they are central to a field, as is the case for Granow et al. (2018). However, in fields such as Communica- tion and Media Psychology we rarely conduct direct replications (Keating & Totzkay, 2019). Consequently, it is unclear how reliable and robust many findings are. Indeed, a large number of findings in the social sciences cannot be replicated successfully (Nosek et al., 2021). Therefore, in the current study we first aimed to revisit and replicate the findings of Granow et al. (2018). In addition, we also extended the original study by elaborating on the relations between binge-watching and self-determination theory. We thereby not only hoped to follow recent calls to subject cen- tral findings to more rigorous tests, we also aimed to demonstrate that a foundational finding on binge-watching can be replicated successfully (Dienlin et al., 2020).

Binge-Watching

Although the phenomenon binge-watching was first defined as early as in 2012 (McNamara, 2012 as cited in Steiner & Xu, 2020), the literature on binge-watching is comparatively young and highly fragmented (Flayelle et al., 2020). However, most researchers agree that binge-

watching differs substantially from traditional forms of series and TV consumption (Jenner, 2017; Panda & Pan- dey, 2017). When binge-watching, individuals continuously consume the same series (e.g., Rubenking et al., 2018), unin- terrupted by commercial breaks or a change of program (Jenner, 2017; Steiner & Xu, 2020), with a certain intensity and a comparatively long reception time (Rubenking & Bracken, 2018). Users can decide which content to con- sume, where, when, and for how long (De Feijter et al., 2016; Jenner, 2017). Often, the reception period is longer than intended. Binge-watching can be planned or unplanned. When unplanned, users often experience a loss of control (De Feijter et al., 2016; Riddle et al., 2018; Rubenking et al., 2018).

Binge-watching usually takes place at home, in most cases alone, and less frequently with a partner or friends (De Feijter et al., 2016; Steiner & Xu, 2020). In most cases, people binge-watch on video-on-demand streaming plat- forms such as Netflix, Amazon Prime, or Disney + (Koch & Beisch, 2020). The number of episodes or the length of the session is seldom determined in advance, and reception depends on the day of the week, the time available, and the type of series being watched (De Feijter et al., 2016, p. 62). In general, video-on-demand consumption is increasing. Whereas in 2017, the year of the original study, only 53% of all Germans watched online video content on a weekly basis, this number rose to 69% in 2020 (Koch & Beisch, 2020). Regarding video-on-demand services in particular, consumption time rose from 40 min in 2018 to 49 min in 2020 (Koch & Beisch, 2020). During the COVID-19 pandemic, during which the data for this study were collected, consumption of video-on-demand services increased even further (Lemenager et al., 2020). However, this does not seem to be an exception, but rather the continuation of an ongoing general trend (Koch & Beisch, 2020).

The main reasons for binge-watching are relaxation and escapism – in other words, mentally switching off and escap- ing from everyday life and its obligations (Panda & Pandey, 2017; Rubenking et al., 2018; Steiner & Xu, 2020; Vaterlaus et al., 2019). Binge-watching has a considerable social com- ponent. Important reasons for binge-watching are recom- mendations from friends and acquaintances, exchanges about what has been watched, and being part of a group (Panda & Pandey, 2017; Rubenking et al., 2018; Shim & Kim, 2018; Steiner & Xu, 2020). Further reasons for and causes of binge-watching are procrastination (Rubenking et al., 2018), entertainment (Panda & Pandey, 2017; Shim & Kim, 2018), being busy (Rubenking et al., 2018),

habits (Rubenking & Bracken, 2018), and perceived autonomy and control (Shim & Kim, 2018; Steiner & Xu, 2020).

Notably, there is still comparatively little quantitative research on the potential impacts of binge-watching on well-being, and many of the existing findings to date stem from qualitative research (e.g., Steiner & Xu, 2020; Vater- laus et al., 2019). For example, Vaterlaus et al. (2019) reported negative relations to health: Physically, in the form of reduced activity and altered eating or sleeping habits; mentally, in the form of additional stress due to procrastina- tion and escapism (Rubenking et al., 2018). Users some- times postpone other activities, such as sports, housework, or university tasks, to binge-watch – both intentionally and unintentionally – and therefore can feel guilty or less relaxed (e.g., Steiner & Xu, 2020; Vaterlaus et al., 2019). Others report more positive outcomes. Groshek et al. (2018), for example, did not find a negative impact of binge-watching on the general health of college students. Rubenking et al. (2018, p. 77) reported that after binge- watching some participants felt good, were more alert than before, and had a sense of accomplishment – especially when they finished a season or an entire series.

Replication and Extension

Granow et al. (2018) were one of the first to investigate quantitatively the relation between binge-watching and well-being and, more specifically, users' recovery experi- ence, vitality, and media enjoyment. Autonomy as well as goal conflicts and feelings of guilt were examined as poten- tial mediators. Granow et al. (2018) thereby approached binge-watching using self-determination theory (SDT; Ryan & Deci, 2000), which makes much sense. However, by focusing on autonomy the authors translated and tested only parts of the theory. For a better understanding of binge-watching, we can now (a) replicate the findings by Granow et al. (2018) and (b) extend the study to test all pre- dictions from SDT. In doing so, we can be more confident regarding the previous finding, while offering a more com- prehensive test of the entire theory. In what follows, we hence first revisit the theory's core tenets and reiterate the hypotheses by Granow et al. (2018), before extending the theory.

SDT is based on the tradition of positive psychology, and it assumes that people want to develop physically and intel- lectually throughout their lives (Ryan & Deci, 2000, p. 68). We all have basic physiological and psychological needs, according to which we set goals and engage in activities. SDT lists three basic psychological needs: competence, autonomy, and relatedness

(Ryan & Deci, 2000). Competence refers to people's need to be effective and efficient in their actions, autonomy addresses voluntarily performing activities or behaviors, and relatedness captures being socially embedded and supported (Ryan & Deci, 2000). SDT understanding of well-being builds on psychological well-being and is oriented toward the eudaimonic well-being of people (Reinecke, 2012). Rather than trying to satisfy only hedonic needs such as positive feelings (Reinecke, 2012), SDT focuses on eudaimonic needs such as self-fulfillment and personal growth.

In general, media use offers numerous opportunities for immediate intrinsic need satisfaction, entertainment, or recreation (Tamborini et al., 2011). At the same time, using media can create conflicts and feelings of guilt that "spoil" users' media enjoyment, which is linked to impaired well- being, both short-term and long-term (Meier et al., 2016). Also regarding binge-watching, Granow et al. (2018) found both positive and negative relations. Binge-watching was related to increased levels of perceived autonomy, which in turn were related to improved media enjoyment and well-being. At the same time, binge-watching was related to goal conflicts, which in turn were related to increased feelings of guilt and reduced well-being, especially concern- ing recovery experience and vitality (Granow et al., 2018). In light of the aforementioned theory and in line with Granow et al. (2018), we therefore posited the following hypotheses[1]:

Hypothesis 1 (H1): Binge-watching is (a) related to more goal conflicts. Goal conflicts are associated with greater feelings of guilt (b). Feelings of guilt (c) are negatively related to users' recovery experience (1), vitality (2), and enjoyment (3).

Hypothesis 2 (H2): Binge-watching is (a) related to more perceived autonomy. Perceived autonomy is
(b) positively associated with users' recovery experience (1), vitality (2), and enjoyment (3).

In adopting SDT, Granow et al. (2018) focused on autonomy. A question that remains unanswered is, when binge- watching, what is the role of competence and relatedness – which are the other two main variables of SDT – and how do they relate to users' well-being? In what follows, we elaborate on why both variables can play a central role when it comes to binge-watching.

We believe competence is a central experience during binge-watching for several reasons. First, noninteractive media use is generally accompanied by a more intense mastery experience, allowing users to feel a sense of com- petence (Reinecke & Eden, 2017). Also, after binge-watching, many users report a sense of achievement and competence, especially after finishing a series (Rubenking & Bracken, 2018). Binge-watching is often intentionally chosen as a leisure activity (Riddle et al., 2018), for example, as a reward or to improve one's mood after a long workday (Rubenking et al., 2018). Thus, being able to make good recreational choices and manage one's mood success- fully might elicit feelings of competence. Finally, binge- watching often occurs in the context of high-quality and complex narratives that usually require close attention (Jenner, 2017). Thus, in consuming this complex content, users demonstrate or at least imply that they are capable of following the complex plot lines, which provides users with feelings of competence.

Hypothesis 3 (H3): Binge-watching is (a) related to more perceived competence, which (b) is positively related to users' recovery experience (1), vitality (2), and enjoyment (3).

As reported earlier, binge-watching has a strong social com- ponent (Rubenking et al., 2018). Users identify with the characters, discuss the series with others, receive recom- mendations, or arrange to watch together (Rubenking et al., 2018). Arguably, there are two major dimensions of relatedness. First, people binge-watch to be able to discuss the plot with friends, to exchange opinions, to speculate about upcoming events, and generally to simply join the conversation (Panda & Pandey, 2017; Shim & Kim, 2018; Steiner & Xu, 2020). The second dimension addresses the content and parasocial interaction. In binge-watching, users can more easily and intensely identify and engage with the series' characters and how they develop in the course of the plot (Shim & Kim, 2018).

Hypothesis 4 (H4): Binge-watching is (a) related to more perceived relatedness, which (b) is positively related to users' recovery experience (1), vitality (2), and enjoyment (3).

Methods

Sample

To answer the hypotheses, we conducted an online cross- sectional survey study. The participants were recruited exclusively online (mainly via Facebook), without receiving financial compensation, and using a snowballing method over a 3-week period in June 2020. In the brief invitation, participants were informed about the survey's topic includ- ing a brief thematic introduction, that no personal data would be stored, and that all information would be anonymized and used for scientific purposes only. Average par- ticipation time was 7 min and 1 s, and median participation time 6 min and 40 s. The language was German. To deter- mine sample size, we conducted an a priori power analysis. Using an alpha level of 5%, a two-tailed test, and a target power of 80%, we aimed to find the smallest significant effect reported in the original study (r = .11). This effect size is in line with the majority of effects in communication, which are mostly small (Rains et al., 2018). Results revealed a necessary sample size of 646. A total of 782 people took part in the survey. We removed participants who were younger than 14, watched series less than once a month, and/or made invalid entries, for example, in the free entry field for age. The final nonrepresentative sample consisted of 668 respondents, who all watched series at least once a month. Participants' ages ranged from 14 to 67 (M = 32.8, Mdn = 31, SD = 9.91), and 81.9% were women. The majority (approx. 72%) were between 20 and 39 years old. Most par- ticipants worked full-time or part-time (about 62%), and only 13% were students. Almost all participants (approx. 93%) stated that they watched series at least several times a week, around 53% of them on a daily basis. A large pro- portion watched predominantly (38%) or exclusively (44.5%) online, with Netflix and Amazon Prime Video being the most frequently used streaming platforms.2

Data Analysis

We analyzed the data using structural equation modeling (SEM). The factorial validity of all measures/scales was tested using confirmatory factor analyses (CFAs). To ascer- tain discriminant validity, we tested all measures together in one single model ("measurement model"), allowing latent factors to covary. All data analyses were performed using the statistical software R. The package lavaan was used for the CFAs and SEMs (Rosseel, 2012). Model fit was assessed using common fit indices (Hu & Bentler, 1999). Our signifi- cance level was 5% and our smallest effect size of interest was r = .10 (i.e., results below .10 were not considered sup- port for a hypothesis, even if statistically significant).3

In general, there is no clear-cut criterion determining the success of a replication (Nosek et al., 2021). Here, we deter- mine replicability by comparing the original study and our replication regarding (a) statistical significance and (b) effect size. In terms of statistical significance, the procedure is as follows. If the original study reports a significant effect, does the replication also find a significant effect in the same direction? Likewise, if the original study does not find a significant effect, is the effect also absent in the replication? Next, when focusing on effect sizes we can compare confidence intervals. Are the confidence intervals of the effects not overlapping, hence significantly different from one another? For example, if the original study finds a small significant effect and the replication a large significant effect, and the confidence intervals of both effects do not overlap, we would not consider the replication successful. It is recommended to compare 83% confidence intervals (Austin & Hux, 2002), ideally coming from sufficiently powered studies. Granow et al. (2018) did not report confidence intervals or standard errors in their paper, and when asked it was technically no longer possible to obtain and provide the results post hoc. Hence, here we compare the point estimates reported by Granow et al. (2018) with the 95% confidence intervals of our study, which is also a possible although a somewhat less precise procedure.

Measurement

The questionnaire was designed on the basis of the items by Granow et al. (2018). In a few cases we needed to change the items (explicated below). In what follows, two example items are presented for each variable (see online supplementary material for all items). All items were answered on a 5-point Likert scale ranging from 1 (= strongly agree) to 5 (= strongly disagree).

We fit the models and the measures as outlined in Granow et al. (2018). However, please note that in several cases the results did not show sufficient model fit or factorial validity. Below, we first report the results of the initial measures as originally fitted by Granow et al. (2018), before we outline how we adapted the measures (later referred to as “adapted measures”).

Binge-Watching. Binge-watching consisted of five items measuring users’ tendency toward intensive series consumption (e.g., “Typically, I watch series straight from first to last episode” or “I like to watch multiple episodes of one series in a single sitting”). Model fit was good, $w2(5) = 21.32$, $p < .001$, CFI = .99, RMSEA = .07, 90% CI [.04, .10], SRMR = .03. Reliability also was good ($\omega = .84$).

Autonomy. Autonomy included four items (e.g., “When watching series, I feel free and self-determined” or “When I watch series, I have a free choice”). Model fit was not good, w2(2) = 58.49, p < .001, CFI = .89, RMSEA = .21, 90% CI [.16, .25], SRMR = .07. Additional factorial analyses revealed two underlying dimensions. The adapted model with two items per factor, with loadings constrained to be equal, showed acceptable fit, w2(3) = 20.44, p < .001, CFI = .97, RMSEA = .09, 90% CI [.06, .13], SRMR = .05. Reliability of the second- order factor also was acceptable (ωL2 = .77).

Goal Conflicts. Goal conflicts were measured with four items (e.g., “When I watch series, I sometimes struggle to stop” or “When I watch series, it sometimes creates conflicts with other things that are important to me”). The unidimensional solution initially suggested did not show sufficient model fit, w2(2) = 119.24, p < .001, CFI = .89, RMSEA = .30, 90% CI [.25, .34], SRMR = .07. Factorial analyses revealed a two-dimensional structure. The fit of the adapted model was good, w2(1) < 0.01, p = .984, CFI = 1.00, RMSEA <.01, 90% CI [< .01, < .01], SRMR < .01. The reliability of the second-order factor likewise was good (ωL2 = .84).

Feelings of Guilt. Feelings of guilt were assessed using a total of five items (e.g., “After watching series, I have a guilty conscience” or “.. . I feel like I must apologize for watching”). As evi- denced by the high RMSEA, model fit was subpar, w2(5) = 104.12, < .001, CFI = .96, RMSEA = .17, 90% CI [.14, .20], SRMR = .04. Modification indices revealed that two items were highly correlated. An adapted model allowing both items to covary showed good fit, w2(4) = 10.91, p = .028, CFI = 1.00, RMSEA = .05, 90% CI [.02, .09], SRMR = .01. Reliability was also good (ω = .90).

Recovery Experience. Recovery experience was measured with eight items (e.g., “When I watch series, I forget about my (work) tasks” or “.. . I relax”). Based on feedback from the pretest, minor linguistic adjustments were made (e.g., “.. . I manage to distance myself from my (work) tasks” instead of “.. . distance myself from work”) to ensure that non-working individuals also feel addressed. The initial model showed poor fit, w2(20)= 1295.81, p < .001, CFI = .55, RMSEA = .31, 90% CI [.29, .32], SRMR = .18. Factorial analyses revealed an underlying two-dimensional structure, with four items per dimension. The fit of the adapted model was good, w2(18) = 78.47, p < .001, CFI = .98, RMSEA = .07, 90% CI [.06, .09], SRMR = .04. The second-order factor of the scale showed poor reliability (ωL2 = .55), though.

Vitality. Vitality comprised a total of 10 items (e.g., active, dynamic, tired), which describe how respondents usually feel after watching series. The items stem from the subscales energy and tiredness of Thayer's (1990) Activation–Deactivation Checklist. In line with Granow et al. (2018), we initially fit a unidimensional model, which did not converge. We then fit a second-order factor model with the aforementioned subdimensions. Model fit was okay, w2(33) = 185.65, p <.001, CFI = .96, RMSEA = .08, 90% CI [.07, .10], SRMR = .04. However, the two subdimensions energy and tired- ness correlated only weakly, also evidenced by the very low reliability of the second order factor (ωL2 = .11). Theoretically, energy seems more closely related to vitality as compared to tiredness, which instead captures lack of vitality. We therefore decided to include only the energy-related aspects in the final analyses. This resulted in a final scale of six items, measuring vitality with a high reliability of ω =.90. Model fit was good, w2(7) = 43.83, p < .001, CFI =.99, RMSEA = .09, 90% CI [.06, .11], SRMR = .02.

Four items (e.g., "Watching series is fun" or "I enjoy watching series") were used to measure enjoyment. Model fit was good, w2(2) = 2.82, p = .245, CFI = 1.00, RMSEA =.02, 90% CI [< .01, .08], SRMR < .01. Reliability was also high (ω = .91).

Competence. Competence was a new variable that we self-designed building on the study by La Guardia et al. (2000). The scale consisted of four items (e.g., "When watching series, I feel that I have made a good choice for my free time"). The factor structure did not hold up in a confirmatory factor analysis. Therefore, we excluded the worst loading item. The adapted model showed improved but still suboptimal fit, w2(1) = 10.49, p = .001, CFI = .92, RMSEA = .12, 90% CI [.06, .19], SRMR = .06. Reliability was also low (ω = .50), limiting the scale's useability.

Six items were self-designed to operationalize relatedness, which consisted of two dimensions with three items each. The first dimension focused on content (e.g., "When watch- ing series, I feel connected to the characters"), the second dimension focused on other people (e.g., "While watching series, I look forward to talking about it with others after- wards"). The model showed good fit, w2(8) = 27.53, p <001, CFI = .98, RMSEA = .06, 90% CI [.04, .09], SRMR =.05. However, reliability was low, and two items also showed insufficient factor loadings. The adapted scale with four items showed good model fit, w2(3) = 10.69, p = .014, CFI = .99, RMSEA = .06, 90% CI [.02, .10], SRMR = .04. However, the reliability of the second

order factor was still low and insufficient (ωL2 = .37). We therefore decided to use a factor-score for the following analyses.

When all adapted measures where fit in one model (mea- surement model), fit was good, w2(739)= 1,487.38, p < .001, CFI = .95, RMSEA = .04, 90% CI [.04, .04], SRMR = .06.

Results

We first looked at the bivariate relations of the variables (see Table 1). Results showed that people who binge- watched more experienced substantially more goal con- flicts, more competence, more autonomy, substantially more relatedness, more recovery experience, and more enjoyment. There were no relevant relations with feelings of guilt and vitality.

Figure 1 compares the means and standard deviations of all variables with the values reported by Granow et al. (2018). Notably, respondents in the original study showed higher levels of autonomy and fewer goal conflicts. The par- ticipants of the replication binge-watched more, which they also enjoyed more.

Replication

We first report the results of our direct replication. Notably, the model we fit, although it followed the procedure out- lined by Granow et al. (2018), showed a different and higher number of degrees of freedom (Granow = 511; repli- cation = 725). We hence assume that Granow et al. (2018) introduced modifications to the model that were not explic- itly reported in the article (e.g., allowing items to covary, or deleting malfunctioning items – both common procedures also applied here). Technically, therefore, we could not reproduce the model tested by Granow et al. (2018).

We then used the adapted measures of the variables as outlined above. The overall fit of the adjusted Model 1 was good, w2(573) = 1,254.82, p < .001, CFI = .95, RMSEA = .04, 90% CI [.04, .05], SRMR = .07. The adjusted model showed a similarly good fit as the model documented by Granow et al. (2018), w2(511) = 1,011.78, p < .001, CFI = .95, RMSEA = .04, SRMR = .05. Likewise, degrees of free- dom were more comparable, suggesting that Granow et al. (2018) tested a similar model. See Figure 2 for a visualiza- tion of the model and the results. To compare both studies, we report the results by Granow et al. (2018) in parentheses.

Results showed that binge-watching was related to increased goal conflicts (β = .47, 95% CI [.40, .55], p < .001). We also found a positive relation between goal con- flicts and guilt: Participants who experienced more conflicts reported more feelings of guilt (β = .19, 95% CI [.10, .28], p< .001). Note that the relationship was substantially weaker than in the original study, where it was β = .47. Next, feel- ings of guilt were not significantly related to recovery expe- rience and vitality. That said, if participants felt more guilty, they experienced slightly less enjoyment (β = -.09, 95% CI [-.01, -.16], p = .024); however, because it is below our SESOI, the effect arguably is too small to matter. Granow et al. (2018) found the opposite pattern. In their study, guilt was related to slightly increased recovery and vitality, but not to enjoyment.

We next analyzed the relation between binge-watching and autonomy. As in the original study, results showed that people who binge-watched more experienced more auton- omy (β = .34, 95% CI [.25, .44], p < .001). Furthermore, as in the original study, respondents who reported more autonomy experienced more recovery (β = .64, 95% CI [.51, .77], p < .001) and more enjoyment (β = .43, 95% CI [.33, .53], p < .001). Unlike in the original study where it was insignificant, experiencing autonomy was related to more vitality (β = .16, 95% CI [.05, .28], p = .007).

When looking at the direct relation between binge-watch- ing and well-being, we found that respondents who binge- watched more also experienced more enjoyment (β = .22, 95% CI [.13, .31], p < .001). No relations with recovery experience and vitality were found. Granow et al. (2018) found the exact same pattern of results.

To test the robustness of the effects, like Granow et al. (2018) in a supplementary analysis we also controlled for gender and age. The model showed acceptable fit, w2(493) = 1,056.54, p < .001, CFI = .95, RMSEA = .04, 90% CI [.04, .04], SRMR = .06. Some additional significant relations with the control variables emerged (e.g., a positive correlation between female gender and autonomy), but the direction of the effects or the significance of the results stayed the same.

Was the replication successful? When looking at significance, the results showed that we replicated eight out of 12 effects, equaling 67%. When looking at effect sizes, the results showed that we replicated five out of 12 effects, or 42%. When combining both criteria, only two effects had the same level of significance and a point estimate within the confidence interval (the relation between binge-watch- ing and recovery and vitality – when controlling for guilt and

autonomy – was nonsignificant in both studies, and effects did not differ in size). For an overview of our results, see Figure 3.

Extension

For our extension of Granow et al. (2018) we fit the same model as before, but did not include goal conflicts and feel- ings of guilt. Instead, we included the SDT variables relat- edness and competence. The model showed good fit to the data, w2(441) = 917.06, p < .001, CFI = .96, RMSEA =.04, 90% CI [.04, .04], SRMR = .06. See Figure 4 for a visualization of the model and the results.

We again tested Hypothesis 2, the relation between autonomy and well-being, but this time also controlling for the relations with relatedness and competence. Three of the four relations were again significant, showing some- what smaller effect sizes. Respondents who binge-watched more also perceived more autonomy (β = .35, 95% CI [.24, .46], p < .001). Experiencing autonomy, in turn, was related substantially to recovery experience (β = .43, 95% CI [.19, .67], p < .001) and enjoyment (β = .27, 95% CI [.04, .50], p = .021). However, there was no significant correlation between autonomy and vitality.

Hypothesis 3 focused on competence. The results showed that people who binge-watched more also experienced more competence (β = .22, 95% CI [.13, .30], p <.001). Similarly, respondents who reported feeling more competent also experienced more recovery (β = .14, 95% CI [.01, .28], p = .031) and vitality (β = .23, 95% CI [.12, .35], p < .001). Feeling competent was not related to enjoyment.

Hypothesis 4 addressed the role of relatedness. Results showed that respondents who binge-watched more felt substantially more related to the characters and coviewers (β =.40, 95% CI [.26, .54], p < .001). In addition, relatedness was a significant and strong predictor of recovery experi- ence (β = .33, 95% CI [.08, .58], p = .010), vitality (β =.34, 95% CI [.12, .55], p = .002) and enjoyment (β = .33, 95% CI [.12, .54], p = .002).

Looking at the direct relations between binge-watching and well-being, now additionally controlling for relatedness and competence, we found the same results as above. Binge-watching was related to slightly increased enjoyment (β = .12, 95% CI [.02, .22], p = .015), but not to recovery experience or vitality.

As in Model 1, we checked the robustness of the effects by including the control variables gender and age. Effects remained largely the same.

Discussion

Binge-watching is a highly relevant current media phenomenon. To date, one of the most important studies on binge-watching was conducted by Granow et al. (2018), which we aimed to replicate here. In a second step, we also included additional variables to test more completely the theory of self-determination in the context of binge- watching.

We replicated the results from Granow et al. (2018) partially. First, we needed to adapt the measures to achieve sufficient factorial validity. Second, we were not able to reproduce the model described in the paper, as our model showed different degrees of freedom. Third, when comparing the results of our new adapted model with the results from Granow et al. (2018), we successfully replicated many of the effects. If we look at significance as a criterion, we replicated 67% of all effects. If looking at effect size, we replicated 42%. In what follows, we elaborate on all three aspects.

First, having to adapt the originally reported measures implies that the quality of the instruments is subpar. From a personal and anecdotal perspective, when conducting confirmatory factor analyses of established and often-used measures, we have often found that factorial validity is insufficient (see also Flake & Fried, 2020). In addition, also our own ad hoc scales on relatedness and competence were suboptimal. Together, the measurement problems reduce the reliability and robustness of our results. We hence emphasize that we all need to invest more time and resources in developing high-quality scales.

Second, not being able to reproduce the originally reported model is surely a limitation. Additional modifications were implemented but not reported in the paper.4 Notably, modifications are not a problem per se, if reported transparently, and if executed only to improve the measurement model but not to achieve statistical significance. Our main take-away is that we should share all information necessary, for example, as supplementary material, so that others can reproduce the study exactly.

Third, from an epistemic perspective, we replicated the finding that binge-watching is associated with heightened conflicts and increased autonomy. In both studies, conflicts were related to more feelings of guilt. The relations between autonomy and well-being were largely

the same: In both Granow et al. (2018) and our replication, respondents who experienced more autonomy reported more recovery and enjoyment. The direct relation between binge-watching and well-being was also comparable in both studies. In conclusion, the relations between binge-watch- ing and well-being, as well as between autonomy and well-being, were largely identical. When looking at the overall potential effect of binge-watching on well-being, which is captured by the bivariate relation of the variables (Rohrer, 2018), we found that people who binged-watched more also reported higher levels of recovery experience and enjoyment (see Table 1). People who binge-watched more than others experienced more guilt. Together, our results imply that we can be increasingly confident that binge- watching has positive relations with both autonomy and some indicators of well-being, but that it is also associated with increased goal conflicts.

We also found some interesting differences. For example, the relation between feelings of guilt and well-being was different. Whereas in Granow et al. (2018) guilt was related to less vitality and less recovery experience but not to less enjoyment, we found the opposite pattern. Future research might hence elaborate on the exact role of feelings of guilt. Notably, effects for guilt were small in both studies, suggesting the need for larger samples to detect them reliably. In addition, whereas Granow et al. (2018) did not find that autonomy was positively related to enjoyment, we found a significant effect. Together, and from a normative perspec- tive, our results are perhaps somewhat more "positive." The relations between binge-watching and conflict were smaller, conflict was more weakly related to guilt, and binge-watch- ing was associated more strongly with autonomy.

How can we explain the divergence? In general, there are four possible explanations. The initial study was technically wrong; the replication was technically wrong; both were technically right, but the underlying phenomenon has changed in the meantime; or both were technically right, but the divergences come from different underlying systemic aspects (also known as "hidden moderators," such as study site, quality, procedure, or sample) or random aspects (such as sampling error or noise). We can only speculate as to what is the right answer in this particular case. When com- paring the means of the variables in both studies it becomes clear that respondents differed. For example, in our replication people binge-watched more, while at the same time they also enjoyed it more. The sample of our replication was a bit older, consisted of more women and considerably fewer students. In addition, the phenomenon of binge- watching

and online-streaming is gaining in importance and relevance. More and perhaps different people are binge-watching today. However, if more and more people binge-watch, the behavior becomes more and more normalized. In the original study, perceived levels of autonomy when binge-watching were much higher than those of our replication, which could potentially suggest that people become increasingly accustomed to the service, and that it has already lost some of its appeal. Hence, the phenomenon indeed might have changed, potentially explain- ing the divergence – while the general pattern still remained the same. In general, it is still an open question as to how much replicability we should expect in a world that is rapidly evolving and changing (Pettigrew, 2018), especially in the context of relatively new media phenom- ena such as binge-watching. Future research should elaborate on this question both from a theoretical and empirical perspective.

In analyzing additional variables from SDT, our second aim was to extend the study by Granow et al. (2018). Results showed that binge-watching was related to increased experiences of relatedness and competence. Respondents who related more to characters and coviewers and who felt more competent also reported higher levels of well-being. These results additionally emphasize that binge- watching is associated with many positive experiences, such as emotional recovery, enjoyment, and vitality. General levels of experienced guilt were low (1.53 on a 5-point scale), whereas levels of enjoyment were high (4.70 on a 5-point scale). Assuming that binge-watching has now become normalized, this is not surprising. The results imply that the somewhat negative societal picture of binge-watch- ing is therefore incomplete.

We would especially like to highlight the pivotal role of relatedness. Engaging in binge-watching comes with substantially increased levels of feeling connected to others – at least, while binge-watching (see Limitations below). The effect size was medium to large ($\beta = .40$), which is a rare finding in the social sciences. In turn, feeling related to both content and coviewers is associated with substantially increased levels of well-being – and again, all effects were of at least medium size (i.e., $\beta > .30$). So far, positive relations between binge-watching and relatedness were found mostly in qualitative research (Panda & Pandey, 2017; Rubenking et al., 2018; Shim & Kim, 2018). There- fore, although binge-watching often takes place at home alone, the experience does feel social. It is possible to exchange with others about recent episodes, making binge-watching a fundamentally social experience with many positive psychological associations.

Limitations

The data were obtained in a cross-sectional survey study, and we controlled for only a handful of variables. As a result, we can make claims concerning relationships, but not causality. It is likely that binge-watching actually increases experienced well-being, but it could also be a selection effect. To make informed statements regarding causality, future research could, for example, try implementing all relevant mediators / control variables (which is hardly feasible), or conduct experiments where each relation is manipulated individually (Rohrer et al., 2021).

Although our study was adequately powered, we recom- mend that future research should run power analyses on the basis of a smallest effect size of interest (e.g., small standardized effects such as r = .10), aiming for a power of 95%. Our study had a chance of finding actually existing effects with a size of r = .10 in 73.5% of all cases, which is decent but can be optimized.

As mentioned earlier, perhaps the biggest weakness of this study is the quality of measurement. We needed to adapt many factors to achieve sufficient factorial validity. In the end, the final model showed good fit, but future research should invest more time in establishing validated scales that are tested in standalone studies. Above all, competence and autonomy were not operationalized with sufficient precision. Future research should revisit the variables and their theoretical underpinnings, to develop a more appropriate scale that clearly distinguishes competence from autonomy. Also note that our measures of competence, relatedness, and autonomy were focused on the binge-watching experience (e.g., “When watching series, I feel free and self-determined”). As a result, our findings do not imply that binge-watching is related to increased feelings of autonomy in general.

Finally, the sampling of respondents can be improved. The current sample had more female participants (82%), who were on average young (72% between 20 and 39 years). To judge better the relationships in the population and compare better results of separate studies, a representative sample would be required.

Conclusion

In this study we were able to partially replicate the findings by Granow et al. (2018). We found that although binge- watching can be linked to negative aspects such as guilt, there are also many positive psychological experiences, including increased relatedness or improved well-

being. Together, our results help explain the widespread adoption of this new and increasingly popular media phenomenon.

References

Austin, P. C., & Hux, J. E. (2002). A brief note on overlapping confidence intervals. Journal of Vascular Surgery, 36(1), 194–195. https://doi.org/10.1067/mva.2002.125015

Birch, J. (2019, June 3). How binge-watching is hazardous to your health. Washington Post,. https://www.washingtonpost.com/ lifestyle/wellness/how-binge-watching-is-hazardous-to-your- health/2019/05/31/03b0d70a-8220-11e9-bce7- 40b4105f7ca0_story.html

De Feijter, D., Khan, V.-J., & van Gisbergen, M. (2016). Confessions of a "guilty" couch potato, understanding and using context to optimize binge-watching behavior. In Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (pp. 59–67). https://doi.org/10.1145/2932206.2932216

Dienlin, T., Johannes, N., Bowman, N. D., Masur, P. K., Engesser, S., Kümpel, A. S., Lukito, J., Bier, L. M., Zhang, R., Johnson,

B. K., Huskey, R., Schneider, F. M., Breuer, J., Parry, D. A.,Vermeulen, I., Fisher, J. T., Banks, J., Weber, R., Ellis, D. A., .. . de Vreese, C. (2020). An agenda for Open Science in commu- nication. In Journal of Communication, Article jqz052. https:// doi.org/10.1093/joc/jqz052

Erdmann, E., & Dienlin, T. (2021). Online Supplementary Material for "Binge-watching, self-determination, and well-being: A partially successful direct replication and extension of Granow et al. (2018)". https://osf.io/6k5rg/

Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. Advances in Methods and Practices in Psychological Science, 3(4), 456–465. https://doi.org/10.1177/2515245920952393

Flayelle, M., Maurage, P., Di Lorenzo, K. R., Vögele, C., Gainsbury, S. M., & Billieux, J. (2020). Binge-watching: What do we know so far? A first systematic review of the evidence. Current Addiction Reports, 7(1), 44–60. https://doi.org/10.1007/ s40429-020-00299-8

Granow, V. C., Reinecke, L., & Ziegele, M. (2018). Binge-watching and psychological well-being: Media use between lack of control and perceived autonomy. Communication Research Reports, 35(5), 392–401. https://doi.org/10.1080/08824096.2018.1525347

Groshek, J., Krongard, S., & Zhang, Y. (2018). Netflix and ill? Emotional and health implications of binge watching streaming TV. Proceedings of the 9th International

Conference on Social Media and Society, 296–300. https://doi.org/10.1145/3217804.3217932

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1–55. 10.1080/10705519909540118

Jenner, M. (2017). Binge-watching: Video-on-demand, quality TV and mainstreaming fandom. International Journal of Cultural Studies, 20(3), 304–320. https://doi.org/10.1177/1367877915606485

Keating, D. M., & Totzkay, D. (2019). We do publish (conceptual) replications (sometimes): Publication trends in communication science, 2007–2016. Annals of the International Communica- tion Association, 1–15. https://doi.org/10.1080/23808985.2019. 1632218

Koch, W., & Beisch, N. (2020). Erneut starke Zuwächse bei Onlinevideo [Online video use continues to rise]. Media Per- spektiven, 9, 482–500.

La Guardia, J. G., Ryan, R. M., Couchman, C. E., & Deci, E. L. (2000). Within-person variation in security of attachment: A self-determination theory perspective on attachment, need fulfillment, and well-being. Journal of Personality and Social Psychology, 79(3), 367–384. https://doi.org/10.1037/00223514.79.3.367

Lemenager, T., Neissner, M., Koopmann, A., Reinhard, I., Georgiadou, E., Müller, A., Kiefer, F., & Hillemacher, T. (2020). COVID-19 lockdown restrictions and online media consumption in Germany. International Journal of Environmental Research and Public Health, 18(1), Article 14. https://doi.org/10.3390/ ijerph18010014

Meier, A., Reinecke, L., & Meltzer, C. E. (2016). “Facebocrastination”? Predictors of using Facebook for procrastination and its effects on students’ well-being. Computers in Human Behavior, 64, 65–76. https://doi.org/10.1016/j.chb.2016.06.011

Merton, R. K. (1974). The sociology of science: Theoretical and empirical investigations. University of Chicago Press.

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Almenberg, A. D., Fidler, F., Hilgard, J., Kline, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L., Schönbrodt, F., & Vazire, S. (2021). Replicability, robustness, and reproducibility in psychological science. PsyArXiv. https:// doi.org/10.31234/osf.io/ksfvq

Panda, S., & Pandey, S. C. (2017). Binge watching and college students: Motivations and outcomes. Young Consumers, 18(4), 425–438. https://doi.org/10.1108/YC-07-2017-00707

Pettigrew, T. F. (2018). The emergence of contextual social psychology. Personality and Social Psychology Bulletin, 44(7), 963–971. https://doi.org/10.1177/0146167218756033

Rains, S. A., Levine, T. R., & Weber, R. (2018). Sixty years of quantitative communication research summarized: Lessons from 149 meta-analyses. Annals of the International Communication Association, 42(2), 105–124. https://doi.org/10.1080/23808985.2018.1446350

Reinecke, L. (2012). Mediennutzung und Well-Being: Zum posi- tiven Wirkungspotenzial interaktiver und nicht interaktiver Unterhaltungsmedien [Media use and well-being: The potential of positive effects of interactive and non-interactive entertain- ment formats]. In L. Reinecke & S. Trepte (Eds.), Unterhaltung in neuen Medien. Perspektiven zur Rezeption und Wirkung von Online-Medien und interaktiven Unterhaltungsformaten (pp. 308–325). Herbert von Halem Verlag.

Reinecke, L., & Eden, A. (2017). Media use and recreation: Media- induced recovery as a link between media exposure and well- being. In L. Reinecke & M. B. Oliver (Eds.), The Routledge handbook of media use and well-being: International perspec- tives on theory and research on positive media effects (pp. 106– 117). Routledge.

Riddle, K., Peebles, A., Davis, C., Xu, F., & Schroeder, E. (2018). The addictive potential of television binge watching: Comparing intentional and unintentional binges. Psychology of Popular Media Culture, 7(4), 589–604. https://doi.org/10.1037/ppm0000167

Rohrer, J. M. (2018). Thinking clearly about correlations and causa- tion: Graphical causal models for observational data. Advances in Methods and Practices in Psychological Science, 24(2), 251524591774562. https://doi.org/10.1177/2515245917745629

Rohrer, J. M., Hünermund, P., Arslan, R. C., & Elson, M. (2021). That's a lot to PROCESS! Pitfalls of Popular Path Models. PsyArXiv. https://doi.org/10.31234/osf.io/paeb7

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. Journal of Statistical Software, 48(2). https://www. jstatsoft.org/v48/i02/paper

Rubenking, B., & Bracken, C. C. (2018). Binge-watching: A suspenseful, emotional, habit. Communication Research Reports, 35(5), 381–391. https://doi.org/10.1080/08824096.2018.1525346

Rubenking, B., Bracken, C. C., Sandoval, J., & Rister, A. (2018). Defining new viewing behaviours: What makes and motivates TV binge-watching? International Journal of Digital Television, 9(1), 69–85. https://doi.org/10.1386/jdtv.9.1.69_1

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55(1), 68–78. https://doi. org/10.1037/0003-066X.55.1.68

Shim, H., & Kim, K. J. (2018). An exploration of the motivations for binge-watching and the role of individual differences. Computers in Human Behavior, 82, 94–100. https://doi.org/10.1016/j. chb.2017.12.032

Steiner, E., & Xu, K. (2020). Binge-watching motivates change: Uses and gratifications of streaming video viewers challenge traditional TV research. Convergence: The International Journal of Research into New Media Technologies, 26(1), 82–101. https://doi.org/10.1177/1354856517750365

Tamborini, R., Grizzard, M., David Bowman, N., Reinecke, L., Lewis,

R. J., & Eden, A. (2011). Media enjoyment as need satisfaction: The contribution of hedonic and nonhedonic needs. Journal of Communication, 61(6), 1025–1042. https://doi.org/10.1111/ j.1460-2466.2011.01593.x

Thayer, R. E. (1990). The biopsychology of mood and arousal.Oxford University Press.

Vaterlaus, J. M., Spruance, L. A., Frantz, K., & Kruger, J. S. (2019). College student television binge watching: Conceptualization, gratifications, and perceived consequences. The Social Science Journal, 56(4), 470–479. https://doi.org/10.1016/j.soscij.2018.10.004

Table 1. Zero-order correlations of all latent variables

	1	2	3	4	5	6	7	8
(1) Binge-watching	–							
(2) Feelings of guilt	−.01	–						
(3) Goal conflicts	.47	.22	–					
(4) Competence	.22	−.32	.10	–				
(5) Autonomy	.35	−.21	.23	.54	–			
(6) Relatedness	.40	.08	.41	.34	.59	–		
(7) Recovery experience	.27	−.15	.30	.48	.69	.61	–	
(8) Vitality	.07	−.09	.13	.28	.20	.33	.22	–
(9) Enjoyment	.36	−.14	.22	.34	.53	.56	.48	.12

Figure 1. Comparison of all variables between Granow et al. (2018) and our replication, using the same operationalization of variables (i.e., not the adapted measures used in the SEMs). Displayed are 83% confidence intervals. Non-overlapping intervals represent significant differences.

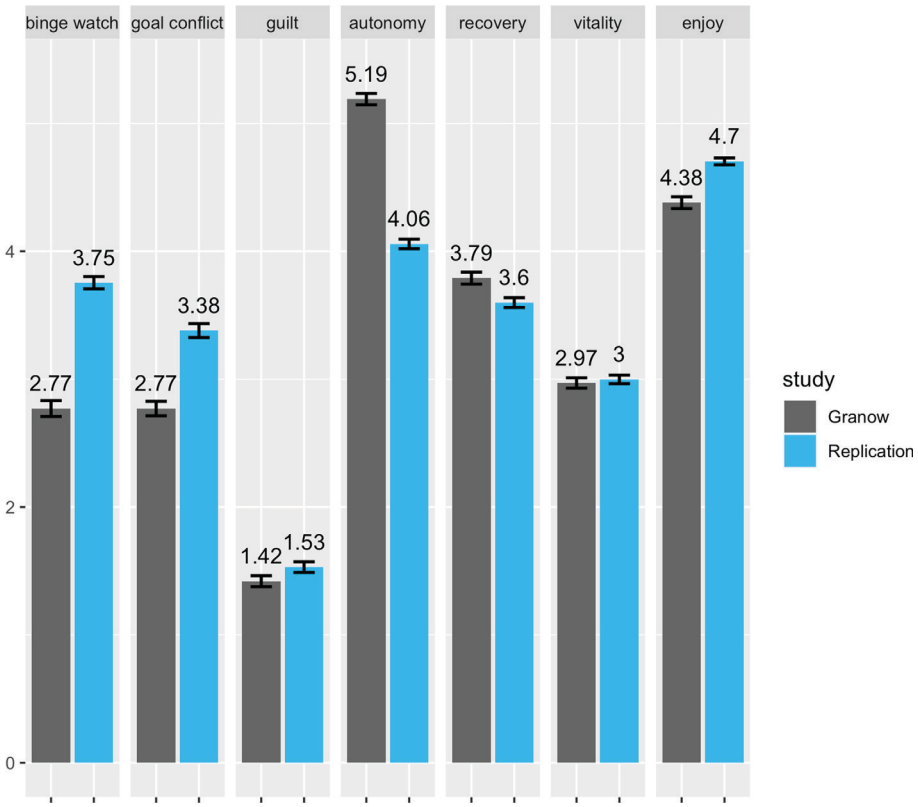

Figure 2. Results from the replication study. Displayed are standardized coefficients. For comparison, results from Granow et al. (2018) are shown in parentheses. To improve readability, the measurement model and the values of the covariances are omitted. *$p < .05$.

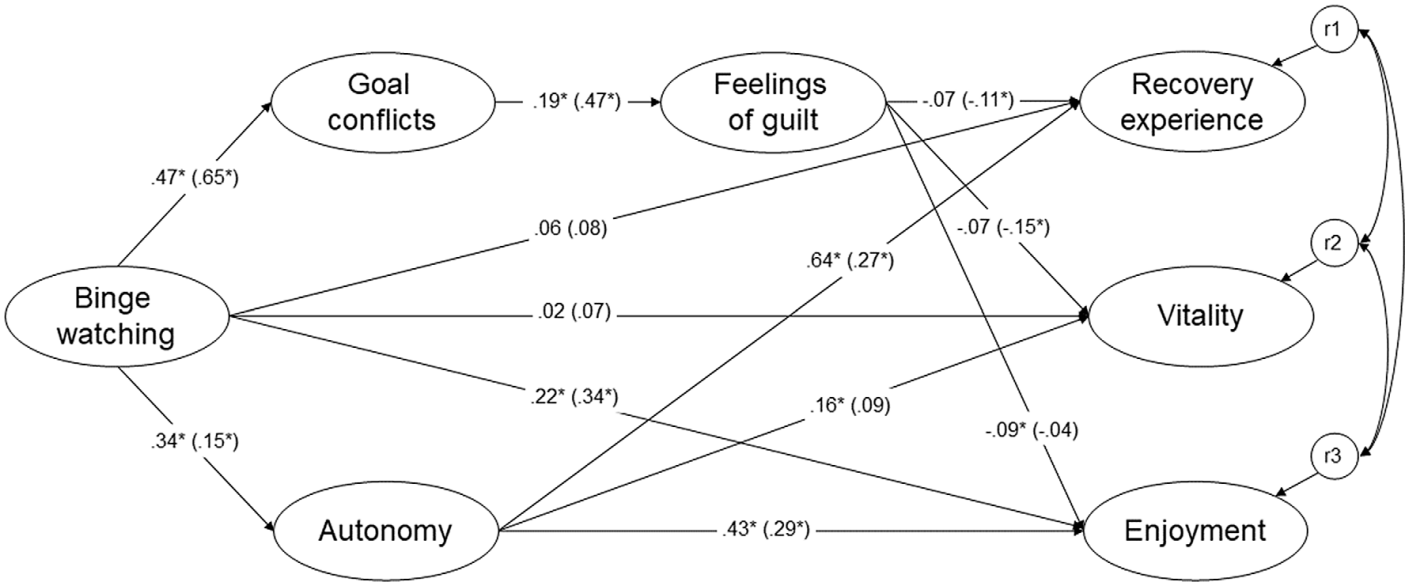

Figure 3. Comparison of Granow et al. (2018) and our replication. Displayed are 95% confidence intervals of standardized effects.