STANDARD PAPER

Writing Yourself Well: Dispositional Self-Reflection Moderates the Effect of a Smartphone App-Based Journaling Intervention on Psychological Wellbeing across Time

Angela MacIsaac1*, Aislin R. Mushquash1 and Christine Wekerle2

1Department of Psychology, Lakehead University, Thunder Bay, Ontario, Canada and 2Department of Pediatrics; Offord Centre for Child Studies, Department of Psychiatry & Behavioural Neurosciences; McMaster University, Hamilton, Ontario, Canada

*Corresponding author: Angela MacIsaac, Department of Psychology, Lakehead University, 955 Oliver Rd, Thunder Bay, ON P7B 5E1, Canada. Email: amacisaa@lakeheadu.ca

(Received 8 March 2022; accepted 3 November 2022)

Abstract

Self-reflection is often viewed positively; paradoxically, however, it is also associated with distress, potentially because of its relationship with rumination. Focusing self-reflection on positive themes may be one way to promote adaptive self-reflection. This study examined whether the disposition to engage in self-reflection motivates use of a journal containing positively focused writing prompts and moderates the benefit gained from it, specifically when rumination is controlled for. For 28 days, participants (N = 152) accessed an app-based mental health intervention containing various features, including the aforementioned journal. Outcomes of self-regulation and psychological wellbeing were assessed, controlling for time spent using other app features. As expected, journaling was associated with improvements in psychological wellbeing but only when baseline self-reflection was average or higher. Journaling was also initially associated with improvements in self-regulation, but this was diminished after controlling for time spent using other app features. Findings suggest self-reflection could be a strength for fostering wellbeing when it is directed in a positive way.

Keywords: self-reflection; journaling; rumination; psychological wellbeing; smartphone app

Introduction

One of the most uniquely human concepts in psychology is that of self-reflection: the ability to think about oneself and inspect and evaluate one’s thoughts, feelings, and behaviours (Grant, Franklin, & Langford, 2002). Like other forms of metacognition, it is part of a collection of higher-order abilities unique to human social evolution (Liotti & Gilbert, 2011). Theoretically, its benefit includes increasing self-knowledge (Fenigstein, Scheier, & Buss, 1975), which in turn promotes self-regulation (Grant, Franklin, & Langford, 2002), personal growth (Harrington & Loffredo, 2010), and adapting to adversity (Borelli, Compare, Snavely, & Decio, 2015). These qualities render the disposition to engage in self-reflection an asset to psychological interventions, including cognitive behavioural interventions, in which an individual’s ability to self-monitor is key to promoting change (Grant, 2001; Sauter, Heyne, Blöte, van Widenfelt, & Westenberg, 2010).

Individuals vary in the degree to which they engage in self-reflection (Grant et al., 2002; Trapnell & Campbell, 1999). While an inclination toward this behaviour should presumably contribute positively to one’s wellbeing and functioning, contemporary examinations of the disposition to engage in self-reflection...

These paradoxical effects may be because self-reflection is related to rumination, a more maladaptive form of self-reflection focused on negative outcomes (Hur, Gaul, & Berenbaum, 2019) and which is associated with psychopathology and a reduced ability to cognitively shift or disengage from information relevant to the self (Beckwé, Deroost, Koster, De Lissnyder, & De Raedt, 2014; du Pont, Rhee, Corley, Hewitt, & Friedman, 2019; Koster, De Lissnyder, Derakshan, & De Raedt, 2011; Norton & Abbott, 2016). Some self-report measures separate rumination from true self-reflection, where the latter is more emotionally neutral, characterised by an intellectual curiosity about the self, and reflected in openness to experience (Trapnell & Campbell, 1999; Treynor, Gonzalez, & Nolen-Hoeksema, 2003). This endeavour has helped clarify that rumination tends to be associated with poor outcomes such as distress and low psychological wellbeing, while self-reflection is associated with higher authenticity, life satisfaction, positive identity, and personal growth (Boyraz & Kuhl, 2015; Harrington & Loffredo, 2010; Luyckx et al., 2007). Despite differential effects on outcomes, however, there is evidence that tendencies to self-reflect and ruminate are related and may co-occur (Takano & Tanno, 2009; Treynor et al., 2003), and as such, any benefit of self-reflection may be contaminated by rumination. Relatedly, self-reflective individuals may make negative judgments about the content of their inner experiences: Harrington et al. (2014) found that while self-reflection was positively correlated with facets of mindfulness of observing, being able to describe, and acting in awareness of internal experiences, it was negatively correlated with the non-judgmental aspect of mindfulness.

**Journaling and Writing Interventions**

Journaling is a behaviour of relevance to dispositional self-reflection. While studies of naturalistic journaling are scant, effects of writing about experiences and feelings have been evaluated in experimental studies. For instance, in the expressive writing paradigm, participants write about a stressful experience with a focus on their deepest thoughts and feelings, resulting in various benefits such as psychological wellbeing and self-regulation (Pascual-Leone, Yeryomenko, Morrison, Arnold, & Kramer, 2016; Pennebaker, Beall, Pennebaker, & Beall, 1986). Of relevance to the current study, out of this literature there has emerged a focus on more positive writing topics (King & Miner, 2000; Lewandowski, 2009; Roepke, Benson, Tsukayama, & Yaden, 2018). For instance, in North, Pai, Hixon, and Holahan (2011), when participants wrote about positive aspects emerging from a stressful situation, outcomes improved in domains of happiness and psychological acceptance. Other positive writing has focused on looking to the future: King (2001) found writing about life goals resulted in improved psychological wellbeing 3 weeks later, and writing about one’s ‘best possible self’ improves positive mood (Heekerens & Eid, 2021). Finally, as opposed to one sole writing prompt, Reiter and Wilz (2016) and Toepfer, Altmann, Risch, and Wilz (2016) used a rotating selection of prompts which focused on themes like finding sources of strength, gratitude, and likeable qualities in oneself, with improvements noted in many domains including rumination. Reflection on these themes may not need to be a time-consuming, in-depth exercise for benefit to be observed: Burton and King (2008) found effects emerged in the domain of health complaints after writing for as little as 2 min over a 2-day period.

Although these writing interventions are associated with beneficial outcomes, they are usually prescribed for a discrete amount of time. There is less appreciation for the fact that many people choose to engage in this kind of writing of their own accord and in a habitual manner in the form of journaling (Duncan & Sheffield, 2008), a sort of ‘self-initiated coping strategy’ (Burt, 1994). This is important considering one meta-analysis found the strongest effects of journaling (about one’s best possible self) were for mood measured directly after writing as opposed to long-term (Heekerens & Eid, 2021).
suggesting writing may work best as an ongoing activity. Like dispositional self-reflection, however, the benefits of journaling may not always manifest in more naturalistic settings. While people tend to keep journals to record, reflect on, and understand emotions (Duncan & Sheffield, 2008; Zammuner, 2001), the wellbeing of journal-keepers seems to diverge from these motives: they are higher in anxiety and distress, while lower in insight (Grant et al., 2002). Interestingly, this is the case despite journal-keepers being higher in self-reflection (Grant et al., 2002). One possibility is that although higher in self-reflection, they may also unintentionally ruminate via journaling (Yukawa, 2008). As such, a journaling intervention with positively focused writing prompts may help guide more adaptive self-reflection.

With respect to the effects of dispositional self-reflection on such an intervention, it is plausible that pairing the disposition to self-reflect with an activity that focuses this behaviour in a positive manner would help promote the beneficial qualities of self-reflection. There is still, however, debate as to whether interventions might be most effective for individuals with a ‘deficit’ in the area or resources available to benefit from them (Toepfer et al., 2016). For instance, those high in rumination seem to benefit from expressive writing regarding prior stressful events (Gortner, Rude, & Pennebaker, 2006; Sloan, Marx, Epstein, & Dobbs, 2008), but did not benefit from positive writing in terms of wellbeing in one study (Toepfer et al., 2016), suggesting the mismatch between rumination and positive writing was not beneficial. No study has evaluated the effect of participants’ self-reflection on such an intervention, particularly the motivation to use such an intervention to begin with, since in prior studies participants have been assigned to use the intervention regardless of their personal characteristics.

The Current Study

While self-reflection, both as a disposition and behaviour via journaling, is thought to be beneficial for overall self-regulation and wellbeing, research suggests that in everyday practice self-reflection may not be practiced adaptively. Since interventions might be most effective for individuals with the characteristics available to benefit from them (Toepfer et al., 2016), individuals with the disposition to self-reflect may be drawn to, and benefit from, an activity focusing on positive self-reflection. The current study aimed to add to the limited research on naturalistic journaling in the context of everyday life by exploring predictors of uptake of a journaling intervention containing positive writing prompts (referred to as the journal or journaling hereafter) as well as the outcomes associated with it. Though journaling is traditionally done via pen and paper, the journal used in the current study existed in the context of a smartphone app. Smartphone use is ubiquitous and many individuals are likely to carry them on their person throughout the day, allowing in-the-moment access to app-based interventions which students in particular have demonstrated an openness to engaging with (Johnson & Kalkbrenner, 2017). We predicted that (1) dispositional self-reflection would motivate greater use of this journal, and (2) dispositional self-reflection would moderate the outcomes associated with its use. Specifically, we theorised those higher in self-reflection would experience more benefit in self-regulation and psychological wellbeing as a result of journaling, only when controlling for rumination.

Method

Study hypotheses were tested in the context of a longitudinal study of the outcomes associated with using the JoyPop™ app (see MacIsaac et al., 2021). JoyPop™ is an app-based mental health intervention containing a journal, among other features.

Participants

Participants consisted of undergraduate students from a Canadian university, recruited through posters, class announcements, and the university’s research study sign-up portal. As part of the larger
study protocol, eligibility criteria included being an undergraduate student, owning an iPhone, and
being able to speak/read/write English fluently. At the time of the study, the app was only available
for the iOS operating system, thus participants were required to have access to a compatible device
in order to participate. Participants received $30 CAD or 2 bonus points towards their grade in a
psychology course for completing the baseline session, $20 CAD for the midpoint session, and $30
CAD for the post-app session (see Procedures section). The number of participants required for
adequate power was assessed using Monte Carlo simulations which is described in MacIsaac
(2020). The final sample of 152 undergraduate students consisted of 115 female participants
(75.66%) and 37 male participants (24.34%), with a mean age of 20.15 (SD = 3.23). 55.92% of parti-
cipants were first-year students, 19.08% were second-year, 14.47% were third-year, and 10.53% were
fourth-year or above.

Measures

Self-Reflection and Insight Scale (SRIS)
The Self-Reflection and Insight Scale (SRIS; Grant et al., 2002) is a 21-item measure of independent
but related constructs of self-reflection and insight. Only the 12-item Self-Reflection subscale that
measures one’s tendency to inspect and evaluate one’s thoughts, feelings, and behaviour was used
in the current study (e.g. ‘I often think about the way I feel about things’). Items are answered on a
6-point scale from ‘strongly disagree’ to ‘strongly agree’. Scores on this measure have previously differ-
entiated those who keep a journal versus those who do not (Grant et al., 2002). In the current sample,
Cronbach’s alpha was 0.90.

Rumination-Reflection Questionnaire (RRQ)
The Rumination-Reflection Questionnaire (RRQ; Trapnell & Campbell, 1999) is a 24-item measure of
rumination and reflection. Only the 12-item Rumination subscale was used. As opposed to its com-
plementary Reflection subscale, which measures reflection motivated by openness to experience, the
Rumination subscale assesses the tendency to engage in self-reflection motivated by neuroticism
(e.g. ‘My attention is often focused on aspects of myself I wish I’d stop thinking about’). Items are
answered on a 5-point scale from ‘1 = Strongly Disagree’ to ‘5 = Strongly Agree’. In the current sample,
Cronbach’s alpha was 0.90.

External Journal Use
Participants were asked the same question as in Grant et al. (2002), with the addition of the part in
brackets: ‘Do you currently keep a journal or diary on a regular basis in which you write about your
thoughts and feelings (outside of an agenda book)?’ This was asked since keeping a journal outside the
study might obviate the need for an additional app-based journal.

Abbreviated Dysregulation Inventory (ADI)
The Abbreviated Dysregulation Inventory (ADI; Mezzich, Tarter, Giancola, & Kirisci, 2001) consists of
30 items measuring difficulties in self-regulation, consisting of the Affect, Cognitive, and Behaviour
subscales. The Affect scale measures self-reported emotion dysregulation including emotional arousal
and control (e.g. ‘I got so frustrated that I often felt like a bomb ready to explode’). The Cognitive
subscales measure self-reported executive functioning capacities which regulate behaviour (e.g. ‘I
developed a plan for all my important goals’). The Behavioural subscale measures self-reported behav-
ioural aspects of dysregulation, such as attention, impulsivity, and hyperactivity (e.g. ‘Little things or
distractions threw me off’). Items are rated on a 4-point scale from ‘0 – Never True’ to ‘3 – Always
True’. As in Mezzich et al. (2001), the three subscales were combined to form a total score.
Cronbach’s alpha values were 0.87, 0.88, and 0.89 for the baseline, midpoint, and post-app timepoints,
respectively.
Ryff Psychological Wellbeing Scale (PWB)

A shortened 21-item version of the original Psychological Wellbeing (PWB; Ryff, 1989) was used to measure psychological wellbeing. We only included the Environmental Mastery, Purpose in Life, and Personal Growth subscales, which were 7 items each. Environmental Mastery measures one’s sense of mastery and competence in managing their environment and capitalising on available opportunities (e.g. ‘In general, I feel I am in charge of the situation in which I live’). Purpose in Life measures one’s belief that life has purpose and is meaningful, as evidenced by the presence of goals and direction (e.g. ‘I have a sense of direction and purpose in life’). Personal Growth measures the extent to which one is continually developing, confronting new problems, and remaining open to new experiences (e.g. ‘For me, life has been a continual process of learning, changing, and growth’). We used these briefer versions of the subscales to avoid respondent burnout while retaining psychometric properties (Ryff, 2014). Items are rated on a 6-point scale from ‘strongly disagree’ to ‘strongly agree’. Springer and Hauser (2006) found the intercorrelations of the three scales used in the current study to be 0.90 and higher, and advised against considering these three scales plus one other scale in the original measure to be distinct. As such, to reduce the number of separate analyses performed, subscales were combined into a total score. Cronbach’s alpha values were 0.89, 0.90, and 0.90 for the baseline, midpoint, and post-app timepoints, respectively.

JoyPop™ App

A comprehensive overview of the JoyPop™ app can be found in MacIsaac et al. (2021) and Mushquash et al. (2021). The app was designed by resilience researchers and youth collaborators and is meant to promote resilience by encouraging the incorporation of the activities it contains into one’s daily routine. It consists of various behavioural features to this end, including a journal. The journal feature contains 28 randomised prompts meant to encourage positive reflection that automatically appear upon opening the feature. Below the prompt there is space for participants to type as much as they wish (there is now also the option to complete entries using voice-to-text technology; this study was conducted prior to the addition of this feature). Journals are saved to the calendar feature. Participants also had access to other app features, including breathing exercises, a mood rating feature, a Tetris-like game, art, and connection to one’s support network and established help lines. Figure 1 depicts the landing page of the app as well as the journal feature with sample writing prompts.
Procedure
Participants attended an in-person baseline orientation session held in group format. During this session, they downloaded the app onto their iPhone and were shown how to use it. Participants then completed all self-report measures. They were then tasked with using the app for 28 days and encouraged to use it at least twice daily, once in the morning and once at night, but were told they were otherwise free to choose how often they used the app and which features they used. This allowed for assessing how often and for how long participants chose to use the journal intervention of their own volition, a strength of the design that mirrored the way naturalistic journaling is a self-initiated habit. Each day during the 28-day period, participants received an email in the morning and at night reminding them to use the app. Participants also attended midpoint (after approximately 14 days) and post-app (after approximately 28 days) sessions to complete the ADI and PWB.

Any time a feature within the app was used, a record of the time and duration was captured. This allowed the focal app variable of journaling, in seconds, to be calculated. Additional variables calculated included the number of days participants used the app across the 28-day study duration henceforth referred to as app days, and the total amount of time, in seconds, spent on all other app features, not including the journal (referred to as other app time). Additional descriptions of these variables can be found in MacIsaac (2020).

Analytic Plan
Data analyses were conducted using Stata. Person-mean imputation was used for missing items within questionnaires. Although this can inflate reliability estimates, the risk of this was lessened as missing items were minimal (Downey & King, 1998), as reported in the Descriptive Statistics section. Furthermore, in the case of a completely missed session, maximum likelihood estimation was used, a preferred method of handling missing data (Dong & Peng, 2013). For each of the analyses described below, a complete description of the model assumptions and diagnostic statistics that were evaluated can be found in MacIsaac (2020), where supplementary analyses are also described.

H1: Self-reflection will predict journal usage
Hypothesis 1 was examined using two analytic methods. First, negative binomial regression was used on the basis that the journaling variable was over-dispersed, such that the variance was greater than the mean, and that the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were lower for this model than competing models (i.e. a Poisson model), indicating better fit. Journaling served as the outcome variable while self-reflection served as the predictor of interest, with rumination, external journal use (yes versus no), and sex included as covariates. Sex was included as a covariate because female participants may be both more likely to journal and more highly self-reflective (Duncan & Sheffield, 2008; Johnson & Wisman, 2013; Zammuner, 2001). App days were included as an exposure variable, which accounts for differences between participants in the amount of opportunity there was to engage in journaling (i.e. the number of days they used the app). With the exception of an outlier with extremely high influence, we retained all other participants in this analysis as opposed to requiring participants to use the feature for some prespecified time, because we were equally as interested in the choice not to journal as the choice to engage with the feature.

Hypothesis 1 was also examined using a hurdle model. Hurdle models are useful for modelling engagement in a behaviour because one function models the binary ‘yes versus no’ choice of whether a behaviour is engaged in at all, and a separate function models how long or how much this behaviour is engaged in if that choice is made. For instance, there may be differences between individuals in not just how long one journals for, but the decision to even engage with the journal to begin with. Conceptualising the ‘choice’ to journal, however, was not straightforward in the current study as very few participants had never opened the journal feature, particularly as we had participants open it when they were introduced to the app, and further, there were data points in which the feature was used for a short period of time (e.g. 2 s) which could indicate an accidental opening of the feature.
Pilot testing was thus conducted with volunteers to assess the average length of time needed to read a journal prompt and write a short paragraph in response to it. A threshold of 30 s was decided upon as constituting the ‘choice’ to engage in a meaningful journal entry, and this constituted the basis of the hurdle model. In the current study, the first function in the hurdle model consisted of a logistic regression predicting whether participants ever chose to engage in a meaningful journal entry, while the second function was a zero-truncated negative binomial regression model predicting the total time spent journaling for those who did. Self-reflection, rumination, external journal use, sex, and app days served as predictors in both parts.

**H2: Journaling and self-reflection will jointly influence outcomes**

Hypothesis 2 was tested using multilevel modelling (MLM). MLM is often used for analysing change over time when both between-subjects and within-subjects (or time-varying) variables exist, particularly as it is robust to missing timepoints or observations. Here, the predictor was journaling, which was totalled for each participant at each timepoint, such that all participants started with a journaling value of 0 at baseline and then had individualised values pertaining to how much they used the journal by the midpoint and post-app sessions. Other app time was totalled in the same manner and served as a control variable to account for the possibility that those who journaled most were also using other app features to a high degree, meaning outcomes could be a result of other app features and not the journal specifically. Between-participant predictors consisted of baseline self-reflection and baseline rumination. A separate MLM was run for each outcome (ADI, PWB).

Models were built hierarchically in three steps following Peugh (2010) [see steps and equations in MacIsaac (2020)]. The first model consisted of journaling as the sole predictor. Then, other app time was added to form the second model. Finally, self-reflection and rumination were added, as well as the interaction of self-reflection with journaling, the primary hypothesised effect of interest. Each successive model was tested for superior fit with a likelihood-ratio test, log likelihood values, and AIC, depending on appropriateness for the model in question. All models were run using full information maximum likelihood estimation and with an independent residual covariance structure specified. The Benjamini-Hochberg correction was used to account for multiple comparisons. In the third model for each outcome, the interaction effect was interpreted in terms of the slope of journaling by creating a Johnson-Neyman plot. This approach involves testing the slope of journaling at each observed value of self-reflection to identify the range of values at which journaling has a statistically significant effect on the outcome. The resulting plot depicts these slopes.

**Results**

**Descriptive Statistics**

Of the 152 participants enrolled, 135 completed their midpoint session (88.82%), and 125 (82.24%) completed their post-app session. In addition to the missing data attributed to these missed sessions, the percentage of missing items across all participants for all attended timepoints (for which person-mean imputation was used) were as follows: 0.22% for the orientation sessions, 0.15% for the midpoint sessions, and 0.22% for the post-app sessions. Means and standard deviations for all self-report measures at each timepoint are presented in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Self-Reflection</th>
<th>Rumination</th>
<th>Psychological Wellbeing</th>
<th>Abbreviated Dysregulation Inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>55.62 (10.94)</td>
<td>3.80 (0.72)</td>
<td>93.78 (15.82)</td>
<td>32.62 (12.87)</td>
</tr>
<tr>
<td><strong>Midpoint</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>95.45 (14.52)</td>
<td></td>
<td></td>
<td>31.93 (12.83)</td>
</tr>
<tr>
<td><strong>Post-App</strong></td>
<td></td>
<td></td>
<td></td>
<td>31.23 (13.41)</td>
</tr>
<tr>
<td></td>
<td>95.57 (15.21)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results also indicated that of the 150 participants who answered the question, 35 individuals (23.33%) indicated that they kept a journal or diary outside of the study. With respect to use of the app, journaling varied widely and was positively skewed, with more values near 0 and fewer at the upper end of the distribution. After removal of an outlier which highly influenced analyses, the minimum value of journaling was 0 and the maximum was 12 h and 41 min; however, the 10th percentile was equal to 25 s of journaling while the 90th percentile was equal to 21 min and 17 s. The median journaling value across the 28 days was 2 min and 17 s, and the median number of times participants opened the journal feature across the 28 days was 4.5 times. This median was similar to art and support network activities (1.76 and 2.72 min, respectively), but lower than the Tetris-like game (15.83 min). It was also lower than the use of the breathing and mood rating features (12.06 and 6.93 min, respectively), but this is expected because these activities were automatically presented to participants when they opened the app. Further description of acceptability of various app features has been documented in Mushquash et al. (2021). The median time spent using the journal feature within a given session was 20 s. Including only ‘meaningful’ journaling sessions of 30 s or higher, this median was 1 min and 24 s. With removal of the outlier with high influence, there were no significant correlations between journaling and age, sex, or use of other app activities.

Does Self-Reflection Predict Use of the Journaling Intervention?

Model output for the negative binomial regression is presented in Table 2. Contrary to the hypothesis that dispositional self-reflection would motivate journaling, self-reflection was not a significant predictor of journaling, nor was rumination or sex. Only external journal use predicted journaling, $\beta = 0.65, p = .025, 95\% \text{ CI} [.08, 1.21]$. In incident rate ratio terms, keeping a journal outside of the app was associated with an increase in the rate of journaling within the app by a factor of 1.91 compared with not keeping a journal. McFadden’s pseudo-$R^2$ remained small at .004; as such, although external journal use had a statistically significant effect on time spent journaling, the overall model explained little of the variability.

Results were similar using the hurdle model, as presented in Table 3. Contrary to predictions, self-reflection was not a significant predictor of the choice to journal in the logistic portion of the model; however, rumination was a significant predictor of the choice to journal, $\beta = 0.56, p = .040, 95\% \text{ CI} [.03, 1.09]$. The odds ratio was 1.75, meaning each unit increase in rumination predicted a 1.75 increase in odds of choosing, at least once, to engage in a meaningful journaling session (of at least 30 s in length) as opposed to never using the journal feature. Similar to the first analysis, external journal use also positively predicted the logistic portion of the model, $\beta = 1.36, p = .003, 95\% \text{ CI} [.47, 2.26]$. The odds ratio for journaling habits was 3.90; that is, keeping a journal outside of the app increased the odds of choosing, at least once, to engage in a meaningful journaling session by a factor of 3.90. The logistic portion of the hurdle model was significant overall at $\chi^2 = 21.98, p = .001$, though the pseudo-$R^2$ for the model was low at .114. In the portion of the model predicting the total time spent journaling once the choice to journal was made, however, no predictors emerged significant.

<table>
<thead>
<tr>
<th>Table 2. Estimates for the Negative Binomial Model Predicting Journaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Self-Reflection</td>
</tr>
<tr>
<td>Rumination</td>
</tr>
<tr>
<td>External Journal Use</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
</tbody>
</table>

* $p < .05$. Note: App days were included in the model as an exposure variable.
Because self-reflection did not predict journaling, an exploratory analysis of the relationship between self-reflection and external journal use was conducted, attempting to replicate Grant et al. (2002). A Welch's t-test indicated those who kept a journal outside of the app were indeed higher in self-reflection at baseline, $t(220.26) = 5.34, p < .001$.

Do Those Higher in Self-Reflection Experience More Benefit from the Journaling Intervention?

Estimates for the three iterations of the psychological wellbeing model are presented in Table 4. *p*-values are significant after controlling for multiple comparisons. There was initially a significant effect of journaling as the sole predictor, $p = .012$, suggesting journaling was positively associated with improved psychological wellbeing improved across the three timepoints. The Wald test for the significance of the fixed effects overall was significant, $\chi^2 (1) = 6.34, p = .012$. Adding other app time to the model did not diminish this effect; journaling remained significant at $p = .043$, while other app time was not a significant predictor, and as such, model fit was not significantly improved. The Wald test was significant at $\chi^2 (2) = 6.43, p = .040$. Finally, upon adding self-reflection and its interaction with journaling, and controlling for rumination, the hypothesis that the benefit of journaling would be augmented by self-reflection was supported: there was a statistically significant interaction, $p = .038$, such that journaling was positively associated with improved psychological wellbeing to the extent that one was self-reflective (while controlling for rumination). A Johnson-Neyman plot of this interaction is presented in Figure 2. The association between journaling and PWB becomes statistically significant ($p < .05$) at self-reflection scores of 46 (i.e. slightly higher than one SD below the mean) and higher; specifically, when self-reflection was at least this high, greater journaling was related to greater PWB. As seen in Figure 2, below this self-reflection score, the association between journaling and PWB does trend negatively; however, none of the slopes were statistically significant ($p > .05$). The Wald test for the significance of the fixed effects overall was significant, $\chi^2 (5) = 48.65, p < .001$. Comparisons of log likelihood and AIC indicated this final model had improved fit compared with the lower-order models.

Table 3. Estimates for the Hurdle Model Predicting Journaling

<table>
<thead>
<tr>
<th></th>
<th>$\beta$ (SE), 95% CIs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Logit regression</strong></td>
<td></td>
</tr>
<tr>
<td>Self-Reflection</td>
<td>$0.001 (0.02), 95% CI [-0.03, 0.03]$</td>
</tr>
<tr>
<td>Rumination</td>
<td>$0.56 (0.27), 95% CI [0.03, 1.09]^{*}$</td>
</tr>
<tr>
<td>External Journal Use</td>
<td>$1.36 (0.46), 95% CI [0.47, 2.26]^{*}$</td>
</tr>
<tr>
<td>Sex</td>
<td>$-0.22 (0.48), 95% CI [-1.16, 0.71]$</td>
</tr>
<tr>
<td>App Days</td>
<td>$0.08 (0.03), 95% CI [0.02, 0.14]^{*}$</td>
</tr>
<tr>
<td>Intercept</td>
<td>$-4.59 (1.64), 95% CI [-7.80, -1.36]$</td>
</tr>
<tr>
<td><strong>Zero-truncated negative binomial regression</strong></td>
<td></td>
</tr>
<tr>
<td>Self-Reflection</td>
<td>$-0.01 (0.02), 95% CI [-0.05, 0.03]$</td>
</tr>
<tr>
<td>Rumination</td>
<td>$-0.05 (0.44), 95% CI [-0.92, 0.81]$</td>
</tr>
<tr>
<td>External Journal Use</td>
<td>$0.33 (0.43), 95% CI [-0.51, 1.17]$</td>
</tr>
<tr>
<td>Sex</td>
<td>$0.26 (0.60), 95% CI [-0.92, 1.44]$</td>
</tr>
<tr>
<td>App Days</td>
<td>$0.07 (0.04), 95% CI [-0.01, 0.14]$</td>
</tr>
<tr>
<td>Intercept</td>
<td>$2.26 (1.87), 95% CI [-1.40, 5.92]$</td>
</tr>
</tbody>
</table>

*p < .05. Note: The logit regression models whether a participant had at least one meaningful journal entry (i.e. for 30 s or more), while the zero-truncated negative binomial regression models the total time spent journaling for those participants who did.
### Table 4. Estimates, Standard Errors, and Confidence Intervals for the Three Iterations of the Psychological Wellbeing MLM Models

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed Effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>94.55 (1.17), 95% CI [92.25, 96.85]*</td>
<td>94.51 (1.18), 95% CI [92.19, 96.82]*</td>
<td>94.52 (1.05), 95% CI [92.46, 96.57]*</td>
</tr>
<tr>
<td>Journaling</td>
<td>.009 (.004), 95% CI [.002, .016]**</td>
<td>.008 (.004), 95% CI [.002, .02]**</td>
<td>.06 (.02), 95% CI [.01, .11]**</td>
</tr>
<tr>
<td>Other App Time</td>
<td>–.00007 (.0002), 95% CI [−.0004, .0005]</td>
<td>–.00005 (.0002), 95% CI [−.0005, .0004]</td>
<td></td>
</tr>
<tr>
<td>Self-Reflection</td>
<td>.33 (.10), 95% CI [.13, 0.52]*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rumination</td>
<td>−8.82 (1.51), 95% CI [−11.77, −5.86]*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journaling × Self-Reflection</td>
<td>.005 (.003), 95% CI [.0003, .01]**</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Variance Components</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>193.44 (23.95), 95% CI [151.77, 246.56]</td>
<td>193.09 (23.93), 95% CI [151.45, 246.19]</td>
<td>146.40 (18.61), 95% CI [114.10, 187.83]</td>
</tr>
<tr>
<td>Residual</td>
<td>38.10 (3.34), 95% CI [32.09, 45.24]</td>
<td>38.13 (3.34), 95% CI [32.11, 45.27]</td>
<td>37.90 (3.33), 95% CI [31.91, 45.02]</td>
</tr>
<tr>
<td>Slope</td>
<td>$1.00 \times 10^{-23}$ (3.45 $\times 10^{-23}$), 95% CI [1.16 $\times 10^{-26}$, 8.60 $\times 10^{-21}$]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p <= .01.

**p <= .05.

Figure 2. Slope of journaling on PWB as a function of self-reflection score. The shaded area represents the confidence interval; the slope is statistically significant when this does not include y = 0 (i.e. at self-reflection = 46 and greater).
Estimates for the three iterations of the self-regulation model are presented in Table 5. There was initially a significant effect of journaling as the sole predictor, \( p = .023 \), suggesting journaling was positively associated with improved self-regulation across the three timepoints. The Wald test for the significance of the fixed effects overall was significant, \( \chi^2 (1) = 5.19, p = .023 \). Upon adding other app time to the model, however, counter to the hypothesis, the effect was diminished and neither predictor was significant; as such, the Wald test was not significant and a likelihood-ratio test indicated model fit was not improved compared with the initial model. Finally, upon adding self-reflection and its interaction with journaling, and controlling for rumination, the main hypothesis was not supported. There was no statistically significant interaction between journaling and self-reflection. There was, however, an effect of self-reflection (conditional on journaling = 0 since an interaction was included, \( p = .001 \)) and main effect of rumination (\( p < .001 \)), such that the Wald test was statistically significant, \( (\chi^2 (5) = 53.09, p < .001) \), and fit was improved compared with lower-order models. This suggests that self-reflection was positively associated with self-regulation overall, while rumination was negatively associated with self-regulation.

### Table 5. Estimates, Standard Errors, and Confidence Intervals for the Three Iterations of the Self-Regulation MLM Models

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed Effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>32.32 (0.99), 95% CI [30.37, 34.27]*</td>
<td>32.38 (1.00), 95% CI [30.43, 34.34]*</td>
<td>32.37 (0.87), 95% CI [30.67, 34.08]*</td>
</tr>
<tr>
<td>Journaling</td>
<td>−.007 (.003), 95% CI [−.014, −.001]**</td>
<td>−.006 (.004), 95% CI [−.014, .001]</td>
<td>−.07 (.04), 95% CI [−.15, .01]</td>
</tr>
<tr>
<td>Other App Time</td>
<td>−.0001 (.0002), 95% CI [−.0005, .0003]</td>
<td>.00004 (.0002), 95% CI [−.0004, .0005]</td>
<td></td>
</tr>
<tr>
<td>Self-Reflection</td>
<td>−.27 (.08), 95% CI [−.43, −.10]*</td>
<td>8.63 (1.24), 95% CI [6.19, 11.06]*</td>
<td></td>
</tr>
<tr>
<td>Ruminatiion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journaling × Self-Reflection</td>
<td>−.0003 (.004), 95% CI [−.008, .007]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Variance Components</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>136.52 (17.24), 95% CI [106.9, 174.85]</td>
<td>135.76 (17.21), 95% CI [105.89, 174.06]</td>
<td>96.51 (12.78), 95% CI [74.44, 125.11]</td>
</tr>
<tr>
<td>Residual</td>
<td>32.55 (2.86), 95% CI [27.41, 38.66]</td>
<td>32.61 (2.87), 95% CI [27.45, 38.75]</td>
<td>31.95 (3.01), 95% CI [26.57, 38.42]</td>
</tr>
<tr>
<td>Slope</td>
<td></td>
<td>.007 (.02), 95% CI [.0001, .49]</td>
<td></td>
</tr>
</tbody>
</table>

*\( p \leq .01 \).  
**\( p \leq .05 \).

Estimates for the three iterations of the self-regulation model are presented in Table 5. There was initially a significant effect of journaling as the sole predictor, \( p = .023 \), suggesting journaling was positively associated with improved self-regulation improved across the three timepoints. The Wald test for the significance of the fixed effects overall was significant, \( \chi^2 (1) = 5.19, p = .023 \). Upon adding other app time to the model, however, counter to the hypothesis, the effect was diminished and neither predictor was significant; as such, the Wald test was not significant and a likelihood-ratio test indicated model fit was not improved compared with the initial model. Finally, upon adding self-reflection and its interaction with journaling, and controlling for rumination, the main hypothesis was not supported. There was no statistically significant interaction between journaling and self-reflection. There was, however, an effect of self-reflection (conditional on journaling = 0 since an interaction was included, \( p = .001 \)) and main effect of rumination (\( p < .001 \)), such that the Wald test was statistically significant, \( (\chi^2 (5) = 53.09, p < .001) \), and fit was improved compared with lower-order models. This suggests that self-reflection was positively associated with self-regulation overall, while rumination was negatively associated with self-regulation.

### Discussion

The current study explored the relevance of self-reflection to a behavioural intervention given that this elusive concept has appeared paradoxically tied to self-awareness and growth yet also distress and rumination. We evaluated whether self-reflection provides the motivation to engage with an app-based journaling intervention and predicts the outcomes associated with it, specifically when statistically controlling for rumination and using prompts geared towards positive self-reflection. We found that on average, uptake of the journal feature was low across a 4-week period. Contrary to the hypothesis that self-reflection would positively predict journaling, there was no association between self-reflection and use of the journal feature of the JoyPop™ app across two analyses, while there was a
positive association between keeping a journal outside of the study and use of the journal feature. Hypotheses were supported with respect to the association between this journaling and psychological wellbeing: although time spent journaling was low on average, it was associated with increased psychological wellbeing even when accounting for the cumulative effect of time spent on other app activities and only for those who had an average or higher tendency to engage in self-reflection. Hypotheses were not supported with respect to self-regulation, with which journaling was not associated once other covariates were added to the model.

That self-reflection did not predict use of the JoyPop™ journal is somewhat contradictory to Grant et al. (2002). Although one strength of our study was refraining from assigning a prescribed amount of journaling, participants were encouraged to use the app in general, which could have led to using the journal to adhere to the study requirements regardless of individual motivation to self-reflect. Another likely explanation is that highly self-reflective individuals were already journaling outside of the app. Indeed, individuals who kept a journal outside of the study were higher in self-reflection, and these same journal-keepers were more likely to use the JoyPop™ journal. This suggests that people who already keep a journal found aspects of the JoyPop™ journal to be acceptable, which is important considering there has been little research on whether technology and pen-and-paper journals are equivalent, despite the availability of journaling features within mental health intervention apps (Alqahtani & Orji, 2020). Furthermore, another predictor of journaling (specifically, the choice to engage in a journaling session for at least 30 s) was rumination, suggesting those with the tendency to engage in negative repetitive thinking can also be encouraged to self-reflect on positive themes. By its nature, rumination it entails thinking about a subject for a lengthy amount of time, though usually in an unhelpful fashion; it is possible that this tendency generalised to mentally engaging with positive topics relevant to the self for a certain amount of time as well.

The most robust finding of the current study was the improvement in psychological wellbeing associated with increased individual time spent using the JoyPop™ journal. This effect persisted when accounting for use of all other app features over time, which supports the specific effect of journaling. This finding is in line with a recent meta-analysis of positive writing, which indicated small to medium effects of positive writing on psychological wellbeing (Carr et al., 2020). The prompts in the journal, touching on themes of growth, mastery, looking to the future, and goals, may have reflected the domains thought to underlie psychological wellbeing. At the same time, however, we found uptake of the journal feature was low for most individuals, rendering the implications of the observed effect tenuous. As suggested below, replications of this effect with a design that encourages increased journal uptake would be helpful. Considering King and Burton (2008) found writing for as little as 2 min had an effect on health, though, the small amount of time many participants in this study spent journaling may not be as detrimental as one would expect.

Those who were not self-reflective did not benefit from journaling in terms of psychological wellbeing, suggesting the prompts alone did not provide the impetus to self-reflect in a manner conducive to positive change. Instead, it seems an intrinsic motivation and ability to self-reflect was needed. Self-reflection has previously been linked to the personal growth aspect of psychological wellbeing (Harrington & Loffredo, 2010), and the two may be theoretically related. The construct of psychological wellbeing captures, among other things, making meaning from one’s experiences, personal growth, and reaching one’s potential (Ryff, 2014). Similarly, self-reflection is purported to increase self-knowledge and insight in the service of psychological growth (Trapnell & Campbell, 1999). In sum, individuals who are high in self-reflection may use positive writing prompts as a resource in shifting the content of one’s thoughts towards positive ideas about the self as opposed to those feeding into anxiety or distress. Future studies might test this idea by asking about participants’ levels of self-focused negative thoughts during the activity or how much the themes of the prompts were novel to participants and captured their attention; alternatively, the text content of entries could be thematically analysed to evaluate the quality of self-reflection.

Despite an initial effect when journaling was the sole predictor in the model, self-regulation did not improve as a function of journaling when controlling for use of the other app features and rumination.
This somewhat contrasts theoretical notions that learning about oneself through self-reflection should allow for improved self-regulation (Grant et al., 2002) and that writing or journaling fosters self-regulation because it provides a sense of control over past and future experiences (Greenberg, Wortman, & Stone, 1996). One possibility stems from Heekerens and Eid’s (2021) finding that the best possible self-positive writing intervention has the biggest effect on mood measured immediately after writing as opposed to long-term. Thus, journaling may promote in-the-moment regulation of emotion not long-lasting enough to be reflected in self-report of self-regulation after a delay. Of course, however, participants in this study journaled for very short durations and increased journaling may produce more robust effects. Alternatively, some theories of self-regulation, such as hypo-egoic self-regulation, suggest that since thinking about our own thoughts and feelings can backfire (e.g. through rumination), sometimes relinquishing attention towards one’s own thoughts can be helpful in achieving self-regulation. For instance, intrinsically engaging tasks that fully direct one’s attention away from oneself, such as playing an instrument or participating in a game, can promote a ‘flow’ state that fosters a sense of enjoyment (Csikszentmihalyi, 1990; Leary, Adams, & Tate, 2006). In this sense, other app activities such as the Tetris-like game or breathing exercises may have been more beneficial for self-regulation, thus reducing the effect of journaling.

Strengths, Limitations, and Future Directions

The current study has many strengths that lend merit to the findings. First, the longitudinal design affords inferences about the directionality of the relationship between journaling and its outcomes, as opposed to past studies of naturalistic journaling which employ cross-sectional designs and say little about whether such observations are a precursor to, or outcome of, journaling (e.g. Grant et al., 2002). A further strength of the current study was allowing participants to decide how much journaling they engaged in. This design represents a compromise between implementing a standardised intervention and considering individual differences that may determine engagement with the intervention. Considering apps are, in essence, self-help interventions in that the user decides when and how often to use it, this design more closely reflects how the app might be used outside of the research context.

Of course, a limitation of the study is the lack of an experimental design. A control group of participants with no app access or access to a journal without positive prompts would have improved the study. Otherwise, the argument could be made that observed changes were due to the natural passage of time. Our use of MLM with journaling as a time-varying predictor specific to the individual lends some protection against this argument (Singer & Willett, 2003), since individuals who did not use the journal or used it rarely serve as a control comparison of sorts. Relatedly, use of other app features was controlled to ensure that the effects were specific to journaling. Controlling for use of the other app features created a high standard for what would constitute real change from journaling as opposed to change occurring due to other forms of intervention in the app. At the same time, it is understandable that when each other activity had its own slightly unique function, the cumulative effect of these different activities might diminish any individual effect of one.

A further potential limitation of the findings is that they may not fully generalise to the form of journaling we are most familiar with; that is, using a physical notebook, especially considering these kinds of journals may not include positive writing prompts. It is unknown whether the effects of journaling would be different without these prompts and whether the negative connotations of self-reflection could manifest. Furthermore, generalisability may be restricted to undergraduate students who own iPhones. One study found some psychological differences between iPhone and Android users (Shaw, Ellis, Kendrick, Ziegler, & Wiseman, 2016), but two larger-scale multi-nation studies found little difference in personality or demographics between them (Götz, Stieger, & Reips, 2017). Nonetheless, results are encouraging for further implementation the JoyPop™ journal and, independent of this, help further elucidate the nature of the construct of self-reflection.

Relatedly, we did not collect or analyse participants’ actual text entries to evaluate the level of engagement with the feature and its prompts; instead, time spent journaling was used as a proxy for
engagement, which comes with its own set of challenges. Time data is associated with a certain level of noise created by participants opening the feature when they did not mean to or potentially continuing to engage with the prompt through other means when the feature was closed. We thus assumed some level of noise across participants and did not exclude participants from analyses except for one outlier. This limitation relates to the hurdle model as well; while we chose 30 s as a cut-off value based on pilot data, it is possible some participants were able to meaningfully engage with the journal for less than this duration. Furthermore, while the short amount of time participants spent journaling might suggest participants did not enjoy or were not interested in the journal itself, a prior qualitative study indicated participants often rated the journal feature as one of the most used or most helpful features, though some suggested having more prompts would be useful (Mushqash et al., 2021). This again suggests time spent journaling is not a perfect proxy for engagement. A more precise estimation of engagement may include analysing entry content; for example, King and Miner (2000) found benefit depended on how many words signalling insight were used, such as ‘understand’ and ‘realise’. At the same time, had participants been informed their responses would be analysed, they may not have been as honest in their entries knowing they would not be kept secret, threatening the validity of the entries.

Given the relative lack of recent research examining naturalistic journaling, and a lack of clarity in prior research on the nature of self-reflection due to its various operationalisations, the current findings may inspire future research. Since many participants used the journal feature for less than 2 min, it would be worthwhile to test the benefits of the journal as a standalone feature of an app, encourage participants to not use other journals while enrolled in the study, or explicitly encourage journaling in the daily reminder emails. If the app is to be integrated into one’s daily routine, however, encouraging people to refrain from going about other habitual activities might not be sustainable. Nonetheless, findings serve as a starting point for further studies of the links between self-reflection, journaling, and wellbeing, particularly those using an experimental design. Researchers could also explore whether rumination might moderate the effect of the JoyPop™ journal, particularly considering our preliminary finding in one analysis that rumination predicted the choice to journal for a meaningful amount of time. As discussed, expressive writing has benefitted those high in rumination (Gortner et al., 2006; Sloan et al., 2008) while in one study, positive writing has not (Toepfer et al., 2016). If rumination does have a negative impact on the outcomes of this intervention, while self-reflection has a positive one, it would be worth considering what this means on a practical level. Would an individual scoring high in both rumination and self-reflection benefit from the intervention or would the opposing effects cancel each other out? Might other writing prompts better address both ruminative tendencies and self-reflection together?

**Conclusion**

Self-reflection is a ubiquitous process, one that is evolutionarily based and necessary for self-awareness and personal growth. The relative benefits and drawbacks of being highly self-reflective have been contested, with true effects confounded by the closely related concept of rumination. Here, although journal uptake was low, preliminary findings suggest that a journaling intervention with positive writing prompts may engage self-reflective processes in a beneficial manner to improve psychological wellbeing. The findings may encourage further research and dissemination of the JoyPop™ journal in a way that encourages uptake, bearing in mind that it may be best suited for individuals whose dispositions are a good fit with the nature of the activity.

**Funding.** This work was supported by the Canadian Institutes of Health Research (grant number TE3-138302).

**Declaration of interest.** None.

**References**


---