EXPRESSIVE WRITING: QUALITY OF LIFE, PRONOUNS, AND WORKING MEMORY

-A PILOT STUDY

Helén Lindgren
I am thankful to Camila Sundvall, former administrator at my (soon to be former) workplace. In spite of changed and very restricted timely resources, she handled all the e-mail contacts, the distribution of codes, and made confidentiality possible. Furthermore her enthusiasm gave me the energy to start the study.

I am thankful to Jeong Jin Yu, Institution of Psychology, University of Umeå, for his supervising and for informing me about the free statistical software JASP. I am also thankful to Tony Qwillbard, Institution of Psychology, University of Umeå, for collecting the AOSPAN results.

And I am grateful to the participants in the current study, who spent quite some time on this experiment in Expressive Writing in Swedish - without any incentives and during a period close to, or even during, Christmas holidays!

Last, but not least, I am grateful to Gunnar, to Albert and Emma, to Amanda and Jessica, and my lovely, little new-born grandchild. Thank you for being the best in my life!
Abstract

For 30 years researchers have studied Expressive Writing (EW) - writing emotionally about a trauma 3-5 times for 15-30 minutes, while controls (CW) objectively write about, most often, time management. Previous research has documented associations, in the trauma condition, between flexible use of pronouns and physical health, or between trauma writing and improvement of working memory (WM). The main aim of the current study was to see if there is a relation between flexible use of first-person pronoun singular (“I”) in EW and improved results on WM test, as well as on quality of life and self-reported illness. In an Internet study participants were recruited through convenience and snowball sampling and randomly divided on conditions. The final sample consisted of 10 participants each in EW and CW: 13 women, 6 men and 1 non-binary, 21-69 years, most of them well educated. Outcome measures were WM test (Automated version of Operation Span Task; AOSPAN), self-reported quality of life (Brunnsviken Brief Quality of Life; BBQ) and self-reported illness (days of symptoms and restricted activities). JASP was used for the statistics. A measure of flexible use of the Swedish pronoun “jag” (“I”) was tested and found to associate with significant improvements in AOSPAN results only; unexpectedly within the whole sample. On condition one significant result was found: EW improved their quality of life. The increase in AOSPAN results is discussed, as well as writing instructions and baseline assessments. The small sample with self-reported health measures and online WM test limited the study. Replications of the results are needed.

Keywords: Expressive writing, time management, working memory, AOSPAN, OSPAN, flexible use of pronouns, first-person pronoun singular, physical health, psychological health, optimism, LOT-R, Quality of Life, BBQ, JASP.
Sammanfattning


Nyckelord: Expressivt skrivande, tidsanvändning, arbetsminne, AOSPAN, OSPAN, flexibel pronomenanvändning, första person pronomen singular, fysisk hälsa, psykologisk hälsa, optimism, LOT-R, Quality of Life, livskvalitet, BBQ, JASP.
Expressive Writing: Quality of Life, pronouns, and working memory?  
- A pilot study

When we are experiencing something difficult, it is better to talk, or write, than just to think (Lyubomirsky, Sousa, & Dickerhof, 2006). When we tell someone what happened, talk to somebody about what makes us stressed, or when we write it down, it seems to lessen the impact on us. More than 30 years ago, trauma researcher professor J. W. Pennebaker and his colleagues started studying what is called Expressive Writing (EW), asking individuals to write expressively, emotionally about their most traumatic and upsetting experiences (e.g., Pennebaker & Beall, 1986; Pennebaker, Colder, & Sharp, 1990). In randomized controlled trials of EW, participants are usually asked to write for 15 to 30 minutes, 3-5 times, in a laboratory setting. The experiment group are instructed to write about their deepest thoughts and feelings concerning a traumatic, personal experience or a stressful life event (like entering college), while the control writing (CW) group are instructed to write in an objective and detailed way about, most often, time management, describing what they have done during specified times before the writing sessions, and sometimes also about what they plan to do on a day to come (e.g., Frattaroli, 2006; Pennebaker & Chung, 2011). Some studies also have included a third condition; positive writing, with the positive writing group instructed to write about, for instance intensely positive experiences or best possible self (King, 2001, King, 2002).

In recent years, participants in EW trials have written not only in laboratory settings, but also at home, on websites (Tatti, 2006), via e-mail (Lu, Zheng, Young, Kagawa-Singer & Loh, 2012), by sending their texts in pre-stamped envelopes (Tavakoli, Lumley, Hijazi, Slavin-Spenny, & Parris, 2009), or without the researchers collecting the texts at all, as in a study examining EW as an expansion to psychotherapy (Graf, Gaudiano, & Geller, 2008).

Physical and psychological health outcomes and overall functioning

EW has been known to be associated with a wide range of physical health measures (number of health clinic visits, self-reported illness, immune function, etc.) or measures of psychological health and well-being (depression, anxiety, stress, PTSD, adjustment to college, quality of life, etc.) (Frattaroli, 2006; Frisina, Borod, & Lepore, 2004; Harris, 2006; Zhou, Wu, An, & Li, 2015). Another kind of outcome is overall functioning (results on exams, school grades, working memory tests, etc.) (Park, Ramirez, & Beilock, 2014; Frattaroli, Thomas, & Lyubomirsky, 2011; Klein & Boals, 2001b).

A more recent and specific physical outcome measure is the time taken to heal punch-biopsy wounds, where participants in the EW condition healed significantly sooner than those in the control group (Weinman, Ebrecht, Scott, Walburn, & Dyson, 2008; Koschwanez et al., 2013; Robinson, Jarrett, Vedhara, & Broadbent, 2017). Lengthy, more or less chronic stress, is known to impair the immune system (Segerstrom & Miller, 2004), while writing about stressful things and experiences seems to have the potential of reducing stress and improving immune function (Petrie, Booth, Pennebaker, Davison, & Thomas, 1995; Petrie, Fontanilla, Thomas, Booth & Pennebaker, 2004; Miller & Cohen, 2001).

But not all studies have shown effects of EW. There are several studies with null effect (e.g., Sohl, Dietrich, Wallstone, & Ridner, 2017; Koschwanez et al., 2017; Niles, Byrne Halton, Mulvenna, Lieberman, & Stanton, 2013; Rivkin, Gustafson, Weingarten, & Chin, 2006; Kloss & Lisman, 2002), and some studies even found a negative effect on the experiment group. Lumley (2004) reported that participants with high scores on alexithymia or aspects of alexithymia did not benefit or even had a small increase in infection symptoms, pain, depression, and anxiety. Also, in a recent study, those in experiment condition, who
scored low on emotional expressiveness, reported increased anxiety at follow up, compared to those in the control condition (Niles et al., 2013).

In an extensive and comprehensive meta-analysis of 146 randomized studies (both published and unpublished) in experimental disclosure, with a median of 60 participants per study, Frattaroli (2006) reported an r-effect size of .075, which the author interpreted as a very small, but in the context yet meaningful effect.

**How come Expressive Writing (sometimes) works?**

No one really knows how EW works (i.e., when it works). A number of theories have been suggested by researchers in EW (e.g., disinhibition theory: writing about something one has never told anyone before; Pennebaker, 1985), self-regulation theory (Cameron & Nicholls, 1998), cognitive-processing theory (e.g., Pennebaker, 1993), social integration model (Pennebaker & Graybeal, 2001), and exposure theory (e.g., Sloan & Marx, 2004).

A theory, akin to the self-regulation theory, is the self-affirmation theory. Self-affirmation is seen as a way to buffer and protect against stress, by strengthening a person’s ability to process difficult thoughts and feelings (Steele, 1988). Creswell et al. (2007) define self-affirmation as “a positive reflection on a valued self-domain” (p. 240), which concerns the current situation, such as e.g., relations, spirituality, self-concept. When coding for self-affirmation units in expressively written texts, a mediating effect between EW and health effects was observed (Creswell et al., 2007).

Another hypothesis, similar to cognitive-processing theory, concerns the importance of creating a narrative (Kennedy-Moore & Watson, 2001; Smyth, True, & Souto, 2001; Klest & Freyd, 2007), which underlines the differences in relating a story and creating a story, as it integrates bits and parts into a holding frame; a narrative. Park, Ayduk, and Kross (2016) looked at the phenomenon of self-distance as a part, and measure, of cognitive processing, and also as a central ingredient in constructing a narrative, since the “I” tells a story about the “me” – a story that makes meaning of the personal experiences. Park et al. (2016) found associations between four variables: a) expressive writing, b) self-distance, c) damped emotional reactivity, and d) less physical symptoms, and described how path-analyses of the data indicated a path from expressive writing, which leads to self-distance, which dampens emotional reactivity, which enhances physical health.

In the meta-analysis of Frattaroli (2006), exposure theory received the greatest support. The hypothesis is that, by habituation, the repeated writing about a traumatic event reduces the impact of an event. So it might be due to the repeated exposure, to the self-distance, to the self-affirmation, to the creation of a story, to a combination of those variables or due to something else, that makes the difference. Pennebaker and Chung (2011) noted the lack of a single and easy explanation of the phenomenon of EW. However, to write, you need a language. Are there any certain aspects of language that could shed light on, or even predict, the effects of EW?

**Language and linguistic markers**

While thinking and keeping our thoughts inside our heads, we are more prone to ruminate (Lyubomirsky, 2006), or more specific, to brood (Nolen-Hoeksema, Wisco & Lyubomirsky, 2008). Putting distress into words seems to facilitate the transforming of the rumination into reflective pondering and adaptive problem solving (Sloan, Marx, Epstein, & Dobbs, 2008). Guastella and Dadds (2006) found that those in the trauma writing condition instructed to see if they could find other ways to think about the incident, seemed to benefit
from, as Larsen and Christenfeld (2011) phrase it; “the ability to relive it in a new way” (p. 145), that is cognitive processing.

When studying changes in thinking, cognitive words are central. Earlier studies have found that an increase of cognitive words (that is causal words: e.g., because, hence, effect and insight words: e.g., think, consider, know) from first to last day in the experiment condition was related to health outcomes (e.g., Pennebaker & Francis, 1996). Furthermore Klein and Boals (2001b), replicated by Kellogg, Mertz, & Morgan (2010), noticed that improvements in working memory (WM) in the trauma condition was associated with an increased use of cognitive words. An increase of cognitive words in the experiment group was also found to connect to improved immune function (Rivkin et al., 2006). Hence, the way we think seem to be associated with both WM and immune function and even with the heart. In a review of cardiovascular recovery from stressful events, Larsen and Christenfeld (2011) made a distinction between cognitive distancing (i.e., distraction from stressor: focusing on something else instead of the stressful experience) and cognitive restructuring (i.e., altering the way of thinking about the stressor), suggesting cognitive flexibility (the ability to vary how one thinks about a traumatic event) to be an important theme in cognitive restructuring. Changing the way of thinking about a stressful event or situation seems helpful. Maybe cognitive flexibility is mirrored by pronoun flexibility?

Campbell and Pennebaker (2003) found a correlation between a flexible use of pronouns in expressively written texts and fewer visits to physicians. This association was not found in the control group. Among the different pronouns the first-person singular pronoun (e.g., I, me, my) is of certain importance, as it reflects depression (e.g., Edwards & Holtzman, 2017; Rude, Gortner, & Pennebaker, 2004), or, according to Tackman et al. (in press), negative emotionality, or even suicide (Stirman & Pennebaker, 2001). Andersson and Conley (2012) found that people with high levels of intrusive thinking gained more when instructed to write in a third-person perspective instead of first-person. But at the same time, especially for persons high in emotional processing, the self-focusing effect of first-person singular pronoun (Dunnack & Park, 2009) and the in-text engagement that follows (Andersson & Conley, 2012) was found to be important in processing stressful life events (Maddalena, Saxey-Reese, & Lester Barnes, 2014). Thus, depending on, for instance, levels of emotional processing, the use of first-person perspective can give different consequences.

As a bridge between pronouns and cognition, Dunnack and Park (2009) found that an increased use of the personal pronoun “I” from first to last text, correlated with an increased use of cognitive words and with better psychological adjustment (they did not examine physical health or WM). Park et al. (2016), on the other hand, found that a decreased use of first-person singular pronouns from first to last text predicted self-distance. The first-person singular pronouns seem to be crucial, but somewhat unclear how.

Two more categories of linguistic markers deserve to be mentioned: one is positive and negative emotional words (e.g., Pennebaker & Francis, 1996) and the other is verb tense and its effects on emotional writing (Tausczik & Pennebaker, 2010) and well-being (Smorti, Pananti, & Rizzo, 2010). Of the four categories of linguistic markers mentioned (pronouns, cognitive words, emotional words and verb tense), the focus in the current study will be on pronouns, more specifically the first-person singular pronoun “jag” (in English “I”). But first we will take a closer look at one of the outcome measures in EW: working memory.
Working memory

Working memory (WM) can be described as “short-term maintenance of information in the absence of sensory input” (Eriksson, Vogel, Lansner, Bergström & Nyberg, 2015, p. 33), or the ability to “keep things in mind while performing complex tasks, such as reasoning, comprehension and learning” (Baddeley, 2010, p. R136). In daily life the WM capacity affects the ability to read, to do arithmetic, to follow a plan, to control responses, etc.

The model of WM function most referred to in neuropsychology, is the one of Baddeley & Hitch (1974), completed by Baddeley (2000) (Gruzka & Orzechowski, 2016). To a three folded first model consisting of a central executive, a phonological loop, and a visuo-spatial sketchpad (Baddeley & Hitch, 1974), Baddeley (2000) added a fourth component; an episodic buffer. As can be seen in Figure 1, Baddeley (2017) describes a hierarchical model where the central executive with its limited attention capacity is focusing on both external perception (stimuli from the surrounding) and internal executive control (e.g., response control). Under the central executive one finds the episodic buffer. The episodic buffer is considered an interface which connects the parts of WM; functioning as a link between the central executive, and the phonological loop and the visuo-spatial sketchpad, respectively. The episodic buffer is seen as a passive storing system, a kind of screen where units of information are accessible to consciousness; information that can be used to plan what to do.
next. Furthermore the episodic buffer is considered an interface between WM and perception, WM and visual and verbal semantics, and WM and long-term memory.

The WM capacity of holding things in mind is stable over time and considered to be a central cognitive trait in a human being (Kane & Engle, 2002). WM capacity as a whole (i.e., both the ability to temporary store and keep things in mind and the functions of attention control: controlling other memory processes, planning what to pay attention to, etc.) is stable and strongly associated with general cognitive ability (Rönnlund, Sundström, & Nilsson, 2015). The WM span (e.g., Turner & Engle, 1989), i.e., how many units a person can remember, in spite of distracting tasks, was shown to be highly correlated to, and able to predict, performances on other cognitive abilities, like prose comprehension, learning, academic performance, abstract reasoning, etc. (Baddeley, 2010; Baddeley, 2017). Turner and Engle’s (1989) test of WM span, Operation Span Task (OSPA) has been used in studies of EW and WM.

Expressive Writing, Stress and Working Memory?

Klein & Boals (2001a) found that the participants’ self-reported life stress predicted their outcomes on WM scores. The authors concluded that life stress creates unwanted thoughts, which demands some of the limited resources needed to perform the WM tasks. Also an alternative explanation was considered; that low WM capacity is a risk factor for life stress, due to lack of ability for effective problem solving. Hence, low WM capacity could either be a consequence of stress, or low WM capacity could worsen the effects of stress (Klein & Boals, 2001a). Recently, with a brief EW intervention of only 2 minutes, Banks & Boals (2017) intended to manipulate an increase in mind wandering short before the participants took a WM test. During the test, the researchers examined the participants’ mind wandering and task-unrelated thoughts. The writing intervention did not alter mind wandering or results on WM tests, but when collapsing the groups it was observed that participants with higher scores on impact of life stress had more task-unrelated thoughts and worse results on the WM tests (Banks & Boals, 2017). Thus, there seem to be interactions between life stress and WM capacity.

Some studies of EW have found that writing in the trauma condition correlates with improvement in WM results, measured by Turner and Engle’s (1989) Operation Span Task (OSPA) (Klein & Boals, 2001b; Yogo & Fujihara, 2008; Kellogg, et al., 2010). The samples in these three studies of EW and WM were all homogenous, consisting of students 18-22 years old (Klein & Boals, 2001b; Yogo & Fujihara, 2008; Kellogg, et al., 2010). Klein & Boals (2001b, Experiment 2) assessed WM before the writing intervention and one week, as well as seven weeks after the intervention, while Yogo & Fujihara (2008) made their follow ups one week and five weeks after the intervention. In the trauma conditions in both these studies there were improvements in results for WM capacity, not at the first follow ups, but at the second follow ups after seven and five weeks, respectively. In the positive writing conditions or control conditions there were significantly less improvements of WM. Noteworthy is that although positive writing has given similar health effects as trauma writing (e.g. King, 2001; King, 2002), it has not resulted in improved WM capacity.

Before the writing intervention Klein & Boals (2001b, in Experiment 2) asked all participants to write a brief description of a positive and a negative event, plus when it started and stopped (if it had). In a replication Kellogg et al. (2010) extended the event recall task. At the first day of the experiment, the whole sample was instructed to recall a stressful event and to write detailed and objectively about it for 15 minutes. Directly after the recall, and also on the third and fifth day, the participants were instructed to write in experiment or control condition. At follow up, at the eighth day, that is only three days after intervention, the EW
group had significantly reduced their symptoms of physical illness, and improved their WM scores. One interpretation was that writing expressively might have resulted in a better coping with intrusive thoughts, leading to freed WM capacity (Kellogg et al., 2010).

In a review Piolat & Bannour (2011) states that most of the studies in EW have aimed at proving that EW changes the affective dynamics of the participants, while these authors call for research aiming at explaining how EW changes it. They recommend investigating EW from cognitive aspects, namely: what editorial processes (e.g., planning, writing, and revising) enhance the emotional functioning of the writers, and what aspects of cognitive architecture (e.g., WM capacity and its handling of attentional resources), restricts the emotional functioning of the writers. In brief: WM is partly considered rather stable and connected to general cognitive ability, partly considered possible to alter by interventions lowering stress levels - such as EW. Based on the earlier research cited, a circle with the interacting factors Stress, Working Memory, and Health, could be sketched, and EW is hypothesized to impact one or all of these factors.

Expressive Writing in Swedish

There are few studies of EW in Swedish. Most of them have focused on the themes or therapeutic processes in the texts (Emilsson, Svensk, Olsson, Lindh & Öster, 2012; Kantus, 2009; Ingels & Damber, 2005; Lindgren, 2007), or has been a single case-study (Overgaard, 2015), while others had very small (Lindgren, 2003) or quite large samples (Kallings, 2006), but without significant results on outcome measures. I could not find any other studies, specifically studying the use of linguistic markers in EW in Swedish, but my own (Lindgren, 2003; Lindgren, 2007).

Aim and hypotheses

An increase of cognitive words, as well as a flexible use of pronouns, is associated with outcome measures of decreased illness. Furthermore an increase of cognitive words is related to improvement in WM capacity, but whether a flexible use of pronouns is associated with improvements in WM capacity, has until now not been studied, as far as I know.

The aim of the present study is to combine these two traces of research concerning EW (i.e., pronouns and health, and WM), and to study it in a sample writing in Swedish, with a follow up period of 2 weeks. The short follow up period is primarily due to the restricted time for the present experiment. In addition, Frattaroli (2006) found in her meta-analysis, that a follow up period of less than a month correlated with a bigger effect size for psychological and physical health and for overall functioning. Because of the restricted time schedule of the present study, it will not be possible to examine cognitive words.

The hypotheses of the current study include:

1) At follow up, the experiment group (writing expressively, emotionally about a negative life event or about something they brood on), compared to the control group (writing objectively about time management), will significantly have: a) decreased their self-reported illness, b) increased their quality of life, and c) increased their results on WM test AOSPAN.

2) The changes in outcome measures in the experiment group will be correlated to the degree of flexible use of the first-person singular pronoun “jag” (in English “I”).
Method

Participants

Demographics: age, gender, and educational level. A total of 58 individuals e-mailed the administrator to participate in the study, of whom 42 completed the base-line questionnaire: Self-report 1. Since one participant resigned from the study after writing the first text, this data was erased. Hence, there were data from 41 participants. Mean age was 43 years ($SD = 16.72$), with a span of 17-75 years. The one minor individual only participated in AOSSPAN pre-test and Self-report 1. Of the 41 participants, 76 % were female, 22 % were male and 2 % non-binary. Education was divided into three levels: a) No university studies: 5 %, b) University studies < 3 years: 15 %, and c) University studies for 3 years or more or studying a program which will last for 3 years or more: 80 % of the participants.

Convenience sample. The author presented the study in four classes of students from the institution of Psychology in the University of Umeå, and information about the study was announced on some billboards at the University of Umeå. Through contacts, e-mails were sent to some students at the institution of Psychology in the University of Lund.

Snowball sampling: Due to difficulties getting enough participants, also snowball sampling was used, where the author (i.e., a white woman, former working class, now middle class, born 1962) by mail and through a private event on Facebook, asked if people in her network could help in recruiting participants.

Assignment to condition: After e-mailing the administrator, participants were randomly distributed to either the experiment group, writing emotionally, or to the control group, writing in a non-emotional way. An intention to match gender on groups was initiated, but not fulfilled, since the participant enrollments arrived in an unpredictable and continuous order. The distribution of women, men and non-binary was made through participant names, since no gender identity was required until the first self-report.

Writing instructions: All participants were asked to write for 20 minutes on three occasions, as in the studies of Klein and Boals (2001b), Yogo and Fujihara (2008), and Kellogg et al. (2010). The participants in the experiment condition were instructed to write for 15 minutes about a stressful life event, which still affects them, and also, for the last five minutes, a) to list what they could do to solve problems or manage challenges, or b) to write about how they have dealt with the difficulties, or c) to see if they despite everything, could find any positive aspects in what they were dealing with (Cameron & Nicholls, 1998; Creswell et al., 2007; Lu & Stanton, 2010; Lu et al., 2012). The control condition was asked to write detailed, objectively and without emotions about what they did yesterday, what they had done today and what they would do tomorrow (e.g., Klein & Boals, 2001b, in Experiment 2; Graybeal, Sexton & Pennebaker, 2002). All participants were instructed to write three times (i.e., three days in a row or within a week) (see Appendix).

The spacing of the sessions does not seem to be of importance for grown-ups (Frattaroli, 2006), but if too narrow spacing (e.g., within 1-3 hours) negative emotions directly after writing gets more intense (Chung & Pennebaker, 2008), especially to adolescents (Travagin, Margola, & Revenson, 2015).

Sheese, Brown, & Graziano (2004) found that 20 minutes approximately corresponded to 500 words. In an e-mail study in EW, instead of a time limit, 500 words was used as a limit for the writing. But since the original instruction was to write for a number of minutes (Pennebaker & Beall, 1986), and since 500 words was a mean and people write in individual tempos, the timing instruction was chosen.

Internet study: To get significant results, a minimum of 20 participants per condition is regarded necessary (e.g., Maddalena et al., 2014). Three earlier studies focusing on EW
and WM had approximately 30 participants in each condition (Klein & Boals, 2001b; Yogo & Fujihara, 2008; Kellogg et al., 2010). Because the intention was to reach 30 x 2 participants in the current study, the decision was taken to make it computer-based, having participants write and self-report over the Internet at a place of their own choice (e.g., Lu et al., 2012, Tatti, 2007), and also to take the WM tests over the Internet.

As seen in Table 1, the study started with the WM test, followed by Self-report 1. After these baseline assessments the writing of the three texts began, followed by Self-report 2 (text-related questions). Two weeks after writing the third text, the participants were once again asked to take the WM test, followed by Self-report 3.

<table>
<thead>
<tr>
<th>Time 1:</th>
<th>Time 2:</th>
<th>Time 3:</th>
<th>Time 4:</th>
<th>Time 5:</th>
<th>Time 6:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Physical health,</td>
<td>Three occasions in</td>
<td>Reflections about the</td>
<td></td>
<td>i.e., Self-</td>
</tr>
<tr>
<td></td>
<td>life quality,</td>
<td>Experiment or Control</td>
<td>writing: How</td>
<td></td>
<td>report 1 and</td>
</tr>
<tr>
<td></td>
<td>optimism,</td>
<td>writing</td>
<td>personal? Emotional</td>
<td></td>
<td>some questions</td>
</tr>
<tr>
<td></td>
<td>impact of</td>
<td></td>
<td>theme?</td>
<td></td>
<td>from Self-report</td>
</tr>
<tr>
<td></td>
<td>traumas</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

### Instruments and materials

**Manipulation check.** An after-writing questionnaire of ten questions was used (Chung & Pennebaker, 2008) to check whether the experimental intervention elicited more personal writing, compared to control. There were also questions about general reactions to writing and about the process of writing, e.g., feeling sad or tired between the writing sessions (Likert scale: 1 (*not at all*) to 7 (*very much*)). Moreover there were five questions derived from Lindgren (2007) about the writer’s view of the emotional content in their texts, e.g., “To what degree did you write about experiences of guilt (shame/grief/anger/worry or fear)” A measure of *Emotions Composite* summed the Likert scale scores on the items about guilt, shame, grief, anger and worry/fear. Questions were also posed to what degree: a) the participants wrote about what affected them most emotionally at the moment?, b) they “censored” themselves?, c) their thoughts wandered?, d) they wrote about experiences in childhood?, and d) to what degree it had been negative to participate in the experiment?, etc. Finally there was an open question, where the participants could express their thoughts about their writing and about the experiment.

**Writing time check.** As mentioned, a way to estimate if the participants wrote for the postulated time is to count the words, where 20 minutes approximately corresponds to 500 words (Sheese et al., 2004).

**Physical health/self-reported illness.** The measures of physical health consist of three questions from the subset Health Related Quality of Life from *Nationella Folkhälsoenkätten Hälso på lika villkor*: e.g., “How many days of the last 30 days were your work ability or your daily activities restricted due to bad physical or psychological health?” (the author’s translation to English for this example) (Folkhälsomyndigheten, 2016) and two questions about days with physical and psychological symptoms respectively. The time specific questions were adapted to half the original period, i.e., the last 15 days, since the
study follow up was planned to be two weeks. Response options ranged from 0 - 15 days. The results for the 15 days before the intervention was subtracted from the number of days of symptoms or restriction of activities at follow up, i.e., 15 last days after the third writing session. One question about general health (Likert scale 1 (very bad) to 5 (very good)) and one question about stress (Likert scale 1 (not at all) to 4 (very much)) (Folkhälsomyndigheten, 2016) were included primary to examine baseline differences. Finally there were three questions about number of experienced traumas (i.e., just counting, not mentioning the traumas), the impact of the traumas and the time since the traumas. These trauma-related questions were included to examine baseline differences.

**Psychological health.** The main measure of psychological health in the study was *Brunnsviken Brief Quality of Life (BBQ)* with 12 statements within six areas of life (e.g., Leisure time, Creativity, Friends, Learning, etc.), where the participants are asked to estimate in Likert scales, 0 (do not agree at all) to 4 (completely agree), both the importance of and the satisfaction with the area. The questions are based on Quality of Life Inventory (QOLI; Frisch, Cornell, Villanueva & Retzlaff, 1992) (Lindner, Andersson, Öst & Carlbring. 2013). The intraclass correlation coefficient (ICC) for BBQ test-retest reliability = .89. Optimal cut-off score between a clinical and non-clinical group for BBQ is 52.5 points (sensitivity = 0.79; specificity = 0.71) (Frykheden, 2014).

*Life Orientation Test-Revised (LOT-R):* This version is a Swedish short form (Muhonen & Torkelson, 2005) based on Scheier, Carver & Bridges (1994), with the intention to assess to what degree a person is optimistic. Optimism is considered a disposition, which promotes a person’s health and well-being, affects coping strategies (Scheier et al., 1994) and even moderates immunological reactions (Segerstrom & Miller, 2004). The difference from the version of LOT-R consisting of ten statements is that the four filler statements have been excluded. The Swedish short form of LOT-R contains six statements (e.g., “In uncertain times, I usually expect the best”, “If something can go wrong for me, it will”), to be estimated in Likert scale 1 (completely agree) to 5 (do not agree at all), where three of the items are reversed. The reliability of the Swedish short form was Cronbach’s $a = .81$ (Muhonen & Torkelson, 2005). In the current study, LOT-R was mainly used to analyze group differences at baseline.

**Working Memory.** As mentioned, in earlier studies of EW and WM (Klein & Boals, 2001b; Yogo & Fujihara 2008; Kellogg et al., 2010) OSPAN (Operation Span Task; Turner & Engle, 1989) have been utilized, but since the current study was an Internet study, WM was assessed with OSPAN in an automated mouse driven version: *Automated Operation Span Task (AOSPAN)*, with an automated scoring system. One difference is that OSPAN uses one-syllable words, while AOSPAN uses letters. The AOSPAN task is to remember a series of letters and simultaneously decide whether simpler mathematical equations are correct or not. If the test taker does not answer within a certain time (computed from the mean time needed for the test person to answer the first part of the test), the test moves on, to prevent participants from rehearsing letters instead of answering the true or false-questions (Unsworth, Heitz, Schrock, & Engle, 2005). The correlation between OSPAN and AOSPAN is $r = .45 (p < .01)$ (Unsworth et al., 2005). If multiplied with itself this would give an $R^2$ of approximately .20, which indicates that circa 20 % of the variance in AOSPAN results would be explained by OSPAN results.

Klein & Fiss (1999) found a practice effect in OSPAN, when administrating different versions of the WM task. There are three presentations of AOSPAN, with an internal consistency of .75 to .80 (Redick et al., 2012), but in the current study the same presentation, the same lists of letters, were used at both pre-test and post-test, which probably might increase the practice effect.
The main score of AOSPAN is the Absolute score, where the number of letters in a fully reproduced correct series of letters is counted, with a total of 75 points. There is also a Partial score, which consists of all correct letters in correct positions. The test-retest reliabilities for the automated Absolute score (Correct List Letters) is .77, while for Partial score (Correct Letters) .83. (Unsworth et al., 2005). Redick et al. (2012) recommend the Partial score to be used, due to its better reliability and its sensitiveness in finding individual differences in WM capacity. Though, since earlier studies on WM Capacity and EW have used OSPAN Absolute (Correct List Letters) score (Klein & Boals, 2001b; Yogo & Fujihara 2008; Kellogg et al., 2010), the choice in the current experiment was AOSPAN Absolute score. OSPAN Absolute Score has a test-retest reliability of .88 (Klein & Fiss, 1999), while AOSPAN Absolute score has lower test-retest reliability, namely .77 (Redick et al., 2009). Thus the test-retest relation considering AOSPAN is less stable.

I did not have access to the WM test results until after the collection of data, but then it was possible to see the exact time when the WM tests were taken (and of course also when self-reports were registered and the dates of the text documents).

**Pronouns.** Earlier studies (e.g., Dunnack and Park, 2009) have in their analyses of language above all used Linguistic Inquiry and Word Count (LIWC) (Pennebaker & Francis, 1999), especially for analysis of emotional words and cognitive words, but also for pronouns (Rude et al., 2004; Dunnack & Park, 2009; Park et al., 2016). Analysis of language has also been made using Latent Semantic Analysis (LSA), where similarities in style words were analysed and a semantic space of 19 pronouns were identified (in decreasing order: I, my, it, you, me, she, he, her, we, they, your, him, his, them, our, myself, their, us, its). The use of pronouns in text 1 was compared to text 2, and the use of pronouns in text 2 was compared to text 3. The results of these comparisons were then averaged (Campbell & Pennebaker, 2003).

Neither LIWC, nor LSA for style words, are available in Swedish, so the challenge was to find a way to measure flexible use of pronouns. One way could have been to tag the parts of speech of the personal, possessive, and reflexive pronouns, as Dimitris Kokkinakis, at Språkbanken, Institutionen för svenska språket, Göteborgs universitet, suggested. In that case, the current 19 pronouns, as translated to Swedish, would have been counted in percent of the total amount of words of each day of the three days of writing, and the standard deviations (SD) for the current pronouns in text 1 and text 2, as well as in text 2 and text 3, would have been calculated, summed and averaged.

But an easier way was chosen: The first-person singular pronoun “jag” (in English “I”) was selected, hypothesized to represent a marker of flexible use of pronouns. In a small study (n = 10), Lindgren (2003) earlier hypothesized (encouraged by J.W. Pennebaker, in a personal communication, 2003), that the standard deviation (SD) for each participant’s percentage use of “jag” (“SDjag”) over the three written texts, could be used as a measure of flexible use of pronouns. Lindgren (2003) found no significant results on outcome measures, though. In the current study SDjag once again was used, but also a more similar approach to the LSA analysis: The SD for the percentage of first-person singular pronoun “jag” in text 1 and text 2, as well as in text 2 and text 3, was calculated, summed and averaged; M:SDjagt1t2&t2t3. Also the SDjag for text 1 and 2; SDjagt1t2, respectively SDjag for text 2 and 3; SDjagt2t3, was used as measures. The last measure was the subtraction of the percentage of “jag” (“I”) in text 1 from the percentage of “jag” (“I”) in text 3; Jagt3-t1 (Dunnack & Park, 2009). Hence, there were five measures of first-person singular pronoun “jag” (“I”), hypothesized to represent a flexible use of pronouns in the texts.
Procedure

1) Informed consent, then a secure mail to participants from the administrator with a personal code and a link to AOSSPAN, followed by a self-report sent through Google Formulary, to assess baseline results. The participants were asked to begin with the AOSSPAN and then proceed with the self-report. 2) Instructions for the writing sessions (experiment condition or control condition) was sent by secure mail, with three coded documents attached. The participants were asked to date and write in one of the three attached documents at a time, for 20 minutes, and to write on three occasions within a week, and then reply the mail with the texts attached. 3) Participants received one reminder to send the texts, where it was said that they could choose to write without sending the text, just announcing that they had written. 4) Soon after having sent the three texts, participants received Self-report 2, with questions concerning their reflections about their writing sessions, 5) A secure follow up mail after 2 weeks from the last day of writing, with the links to AOSSPAN and Self-report 3, and if necessary, 6) A reminder to do the follow ups.

Statistical Analyses

The final sample size was small \( (n = 20) \), but at least the two study groups were equal, which is favorable when it comes to the robustness of statistical tests (Boneau, 1960; Curtis et al., 2015). By the notion robustness, statisticians mean a low possibility of an incorrect conclusion. There is a long-lived discussion about: a) the use of parametric versus non-parametric tests, especially with a small sample size (< 30 participants), b) violations of the assumption of equal variance, and c) ordinal scales (as Likert scales), but the explicit conclusion of Norman (2010) is that parametric tests are robust even in these cases. Norman (2010) writes: “A small sample size makes the hurdle higher, but if you´ve cleared it, you´re there (p. 628)”.

The outcome measures were five to the number: AOSSPAN Absolute score (working memory), BBQ (Quality of Life), number of Days with Physical symptoms, number of Days with Psychological symptoms, and number of Days with Restricted activities or work ability due to physical or psychological symptoms. By “Change-results” is meant baseline (pre-test) subtracted from follow up (post-test). Another way to put it is follow up minus baseline (post-test minus pre-test).

The free software JASP version 0.8.5 (JASP Team, 2018) was used for all statistical analyses. Chi-square was used for comparisons of categorical variables on groups. Independent samples student \( t \)-tests were used to assess baseline differences between completers and dropouts, as well as between experiment group and control group within the final sample. Besides from baseline differences and manipulation checks, single Likert scale items were not analysed (Carifio & Perla, 2008). Pearson product-moment correlations was used for correlations among baseline individual differences and the outcome measures. Correlations were also run between variables from the writing intervention and change-results, to identify covariates to be used in the Repeated Measures ANOVA. As in the studies of Klein & Boals (2001b), Yogo & Fujihara (2008), and Kellogg et al. (2010), Repeated Measures ANOVA (time x condition) was used to assess outcome measures, with pre-test and post-test as time, and experiment group and control group as condition. Additionally Median split analyses was used to examine some outcome results more thoroughly.

In few cases in the current study, there were violations of the assumption of equal variances in Independent samples Student \( t \)-test, but since the groups were equal, this was of minor importance (Boneau, 1960).
Ethical considerations

To protect the integrity of the participants, the ”secure mail” (i.e., encrypted e-mail) of my workplace (at an occupational health care) was used. Each participant received a code, administered by the administrator at my workplace, to protect the confidentiality of the participants.

In an informed consent the participants received information about the study, about confidentiality, the right to withdraw, how the data will be used, and that, due to what they would write about, the writing could result in being negatively emotionally affected a while after writing. In the writing instruction for the experimental condition participants were informed that if having a very strong emotional reaction they should stop writing and take a walk or contact a friend, and that some people feel low directly after writing (Pennebaker et al., 1988; Hockemeyer, Smyth, Anderson, & Stone, 1999), and if feeling they would “flip out” by writing about something, they should not write about it (Pennebaker & Evans, 2014).

The writing instruction chosen for the current study is presented with an addition of a choice between self-regulation (to list things that could help them cope with problems) (Cameron & Nicholls, 1998), self-affirmation (Creswell et al., 2007, Steele, 1988) (here: how the person dealt with the difficulties) or cognitive reappraisal/cognitive flexibility (if there had been any positive aspects of the event, like changes being made as a result of the experiences) (Lu and Stanton, 2010). The aim with this amendment was to strengthen the possibility that the writing in the experimental condition would promote positive emotional outcome. Creswell et al. (2007) furthermore noticed that participants who wrote in a more self-affirmative way and with more combinations of cognitive processing and discovery of meaning, were less distressed directly after the writing sessions. Finally an earlier e-mail based study in EW reported no more negative responses in experimental condition than in control condition (Sheese et al., 2004).

Results

Attrition, missing data, and outliers

Attrition. In Table 2 is shown, that of the 58 individuals who e-mailed the administrator, 40 received results on the WM pre-test and 42 individuals filled in Self-report 1. As mentioned one participant resigned, hence this data was erased. Thirty-five participants wrote the first text, while 32 wrote all three texts, of whom 28 e-mailed all three texts. Of the 32 participants who completed the writing task (whether or not they e-mailed their texts), 30 individuals filled in Self-report 2, while 25 received results on the WM post-test and 26 participants filled in Self-report 3. In brief: Of the 58 individuals who signed up, 27 participants fulfilled at least one of the two follow-up-measures, which gives an attrition rate of 53%.

Missing data. Probably due to technical issues, WM results were not complete for three participants (two in Experiment Group and one in Control Group). Furthermore Self-report 3 was lacking for one participant in Control Group. These four participants with incomplete data were excluded. Experiment Group now consisted of 8 women and 5 men, while Control Group included 7 women, 2 men and 1 non-binary.

Outliers. Some of the Change-results for AOSpan Absolute Score (Correct List Letter Score) were extraordinary. When the data was examined through Interquartile Range, three outliers were found, two women and one man, all in Experiment Group. After the exclusion of these three outliers, no more outliers were identified.

The final sample of 20 participants, out of 58 registered, gives an attrition of 66%.
Table 2
Attrition and missing data

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Registered participants</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>58 randomized to conditions</td>
<td>27</td>
</tr>
<tr>
<td>22 (1 missing result)</td>
<td>40 received results on AOSPAN, pre test</td>
<td>18 (1 missing result)</td>
</tr>
<tr>
<td>22: 14 women, 8 men</td>
<td>42 did Self-Report 1</td>
<td>20: 17 women, 2 men, 1 non-binary</td>
</tr>
<tr>
<td>18 (1 resigned from study)</td>
<td>35 wrote text 1</td>
<td>17</td>
</tr>
<tr>
<td>16</td>
<td>32 wrote all 3 texts</td>
<td>16</td>
</tr>
<tr>
<td>14</td>
<td>28 e-mailed all 3 texts (optional)</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>30 did Self-Report 2</td>
<td>15</td>
</tr>
<tr>
<td>13: 8 women, 5 men (2 missing results)</td>
<td>25 received results on AOSPAN, post-test</td>
<td>12: 9 women, 2 men, 1 non-binary</td>
</tr>
<tr>
<td>15: 10 women, 5 men</td>
<td>26 did Self-report 3</td>
<td>11: 8 women, 2 men, 1 non-binary</td>
</tr>
<tr>
<td>15: 10 women, 5 men</td>
<td>27 received results on at least one of baseline measures and one of follow up measures</td>
<td>12: 9 women, 2 men, 1 non-binary</td>
</tr>
<tr>
<td>2 women excluded (1 with both pre &amp; post-test missing)</td>
<td>3 excluded due to missing results on AOSPAN pre-test or post-test</td>
<td>1 woman excluded</td>
</tr>
<tr>
<td>None excluded</td>
<td>1 excluded due to missing results on Self-report 3</td>
<td>1 woman excluded</td>
</tr>
<tr>
<td>2 women, 1 man excluded</td>
<td>3 outliers excluded due to results on AOSPAN Change results</td>
<td>None excluded</td>
</tr>
<tr>
<td>10: 6 women, 4 men</td>
<td>Final sample: 20 participants with complete data</td>
<td>10: 7 women, 2 men, 1 non-binary</td>
</tr>
</tbody>
</table>

**Dropouts:** younger, fewer Days with physical symptoms, and less optimistic

“Completers” are here defined as those who did at least one part of follow up, that is AOSPAN post-test and/or Self-report 3. In an Independent Samples t-test, Completers (n=27) were shown to be older (M = 48.15; SD = 16.05) than Dropouts (n=14) (M = 33; SD = 13.26); t(39) = -3.09, p < .01, d = -1.02. The Completers reported more Days of physical symptoms (M = 4.04; SD = 4.75) compared to Non-completers (M = 0.54; SD = 0.66); t(39) = -2.62, p = .012, d = -.89, and were more optimistic, as measured by LOT-R (M = 22.07; SD = 4.47), compared to Dropouts (M = 18.79; SD = 2.99); t(39) = -2.47, p = .02, d = -.81. A Chi-squared test showed no significant difference on gender for Completers (19 women, 7 men, 1 non-binary) and Dropouts (12 women, 2 men) ($\chi^2 = 1.38, p = .50$). Hence, in comparison with Completers, Dropouts were younger, reported fewer Days of physical symptoms and were less optimistic.
The final sample: more optimistic compared to the others who answered Self-report 1

From the Completers (n = 27) four participants with incomplete results on follow up, and three outliers on AOSPAN Change-results, were excluded. In a baseline analysis of the final sample (n = 20), compared to the rest of the participants who filled in Self-Report 1 (n = 21) a Chi-squared test showed no significant difference on gender (p = .26). Only one significant group difference was found with an Independent samples t-test: The final sample were more optimistic (LOT-R) (M = 22.35, SD = 4.80) in comparison with the rest of those who answered Self-Report 1 (M = 19.62; SD = 3.32); t(39) = -2.13, p = .04, d = .66.

Baseline differences: Control Group reported more days of restricted activities

Experiment Group (n = 10) consisted of 6 women and 4 men, while Control Group (n = 10) consisted of 7 women, 2 men and 1 non-binary. Chi-squared tests showed no significant group difference on gender (p = .42), or education (p = .59). An Independent samples t-test showed no significant differences in age between Experiment group and Control group (ps > .05). The only significant baseline difference between conditions in final sample was Days of restricted activities due to physical or psychological symptoms; t(18) = 2.82, p = .011, d = 1.26. The control group reported significantly more Days of restricted activities (M = 4.80; SD = 4.29) at baseline, compared to the experiment group (M = 0.80; SD = 1.32). In Table 3 descriptive statistics for baseline measures on condition can be seen.

A Pearson’s product-moment correlation was run to test associations between age, education and outcome measures at baseline. Two significant correlations were found, both concerning Days of psychological symptoms, which was negatively related to: a) Age (r = -.51, p = .02), and b) Quality of Life (BBQ) (r = -.52, p = .02). This indicates that younger age and lower score on Quality of Life (BBQ) respectively, were associated with more Days of psychological symptoms, though Age and Quality of Life was not mutually correlated.

The final sample: Manipulation check

Text-related questions. The first ten questions in Self-report 2 were manipulation check questions from Chung & Pennebaker (2008) and significant differences between groups showed that the intervention was successful. In comparison with the control group, the experiment group reported their writing to be more personal, meaningful and emotional. Also they found it more difficult to write, felt sadder and more tired between the writing occasions, and thought more of their subjects of writing between the writing sessions. Furthermore the experiment group reported understanding more of, and thinking differently about, the subjects they wrote about (all ps < .01, d = -2.08 to -2.99). There were also significant differences in opinions about the length of writing. The experiment group reported the time to write being too short (p = .007, d = 1.35). Also for the other text-related questions, there were significant differences between groups, since the experiment group reported writing more emotionally and having new thoughts from writing (all ps < .001, d = 1.35 to -6.57). There were no significant differences between groups concerning to what degree the participants censured themselves, or to what degree their minds wandered during the writing (ps > .05); in both groups participants censured themselves and their thoughts wandered, to some degree.
Table 3
*Descriptive Statistics of Baseline differences between Experiment and Control groups in final sample (n=20).*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>44.50</td>
<td>46.40</td>
<td>5.40</td>
<td>5.80</td>
<td>3.60</td>
<td>4.00</td>
<td>5.10</td>
<td>3.20</td>
<td>4.80</td>
<td>0.80</td>
<td>1.90</td>
<td>2.00</td>
<td>59.50</td>
<td>65.30</td>
</tr>
<tr>
<td>Education</td>
<td>5.20</td>
<td>5.23</td>
<td>0.40</td>
<td>0.36</td>
<td>0.21</td>
<td>0.23</td>
<td>1.37</td>
<td>1.44</td>
<td>1.57</td>
<td>0.58</td>
<td>1.36</td>
<td>0.42</td>
<td>0.28</td>
<td>0.21</td>
</tr>
<tr>
<td>General health pre</td>
<td>27.1</td>
<td>27.3</td>
<td>1.60</td>
<td>1.80</td>
<td>0.44</td>
<td>0.77</td>
<td>18.77</td>
<td>20.84</td>
<td>24.62</td>
<td>3.39</td>
<td>18.40</td>
<td>1.73</td>
<td>0.77</td>
<td>0.44</td>
</tr>
<tr>
<td>Physical days pre</td>
<td>276.0</td>
<td>276.0</td>
<td>534.9</td>
<td>276.0</td>
<td>25.21</td>
<td>23.38</td>
<td>61.96</td>
<td>32.93</td>
<td>1.88</td>
<td>0.93</td>
<td>1.57</td>
<td>5.29</td>
<td>210.4</td>
<td>521.9</td>
</tr>
<tr>
<td>Psychological days pre</td>
<td>17.65</td>
<td>17.65</td>
<td>0.08</td>
<td>0.07</td>
<td>0.91</td>
<td>0.27</td>
<td>1.69</td>
<td>1.91</td>
<td>0.22</td>
<td>0.00</td>
<td>0.57</td>
<td>0.38</td>
<td>-0.56</td>
<td>-0.73</td>
</tr>
<tr>
<td>Restricted days pre</td>
<td>0.34</td>
<td>0.34</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>Stress pre</td>
<td>2.50</td>
<td>2.60</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>BBQ pre</td>
<td>1.90</td>
<td>2.00</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>LOT-R pre</td>
<td>2.00</td>
<td>2.00</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Nr of trauma</td>
<td>2.50</td>
<td>2.60</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>Impact trauma</td>
<td>0.80</td>
<td>1.64</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>Time since trauma</td>
<td>2.50</td>
<td>2.60</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>AOSPAN Absolute pre</td>
<td>38.0</td>
<td>38.0</td>
<td>5.78</td>
<td>5.78</td>
<td>5.78</td>
<td>5.78</td>
<td>5.78</td>
<td>5.78</td>
<td>5.78</td>
<td>5.78</td>
<td>5.78</td>
<td>5.78</td>
<td>5.78</td>
<td>5.78</td>
</tr>
</tbody>
</table>

Mean

Std. Error of Mean

Median

Std. Deviation

Variance

Skewness

Kurtosis

Range

Minimum

Maximum
The order of the parts in the experiment. An analysis of the registered times of WM-tests, self-reports and texts in the final sample, showed that one participant began with Self-report 1 (instead of AOSPAN pre-test), while all other participants performed the parts in the stipulated order.

Time and amount of writing. A word count was made to estimate if the participants wrote for 20 minutes. The mean for all participants who wrote at least one text (n = 34), was 498.60 words (SD = 190, Span: 171-931). Within the final sample (n = 20), the Experiment group wrote significantly more words ($M = 633.70; SD = 211.80$, Span: 232-935), compared to Control ($M = 417.7; SD = 120.90$, Span: 272-587); t(17) = -2.77, p = .01, d = -1.27.

Number of days from intervention to follow up. The time between Text 3 and WM 2 was controlled; $M = 19$ days; SD = 4.46, span 13-27 days. Thus the time to follow up varied and was longer than the intended two weeks.

Correlations between baseline Optimism, Pronoun flexibility, Emotions Composite, Time to follow up, and Outcome measures change-results

As seen in Table 4, Pearson product-moment correlations were run between pronoun flexibility, self-reported emotionality, time to follow up, baseline optimism and the outcome measures, to see which of them to test as covariates in the Repeated Measures ANOVAs.

Optimism. No significant correlation was found on optimism. The mean score on optimism (LOT-R) for the final sample (n = 20) was 22.35 (SD = 4.80), which is close to the Swedish mean score of students: 21.44 (SD = 4.52) (mean age 26 years; $SD = 6.32$) (Muhonen & Torkelson, 2005). Additionally, as in Cameron and Nicholls (1998), a Median split on LOT-R was made, with half the group above the median and half beneath. A Repeated Measures ANOVA (time: pre-test, post-test x condition: median split) showed no significant results for any of the outcome measures (all $ps > .05$). Thus, the level of optimism did not impact the results within the final sample.

First-person singular pronoun. Significant correlations were found between AOSPAN change-results (pre-test score subtracted from post-test score) and the four measures of Standard Deviation for the use of first-person singular pronoun “jag” (in English “I”), with the strongest correlations found for SDjag2t3 ($r = .59$, $p < .01$) and MeanSDjagt1t2t2t3 ($r = .58$, $p < .01$). A flexible use of the first-person singular pronoun “jag” was associated with improvement in AOSPAN Absolute score.

Emotions Composite. Emotions Composite (the sum of the self-reported degree of writing about guilt, shame, grief, anger and worry/fear) was significantly associated with BBQ change scores ($r = .60$, $p < .01$). To write about negative emotions seemed to correlate to a notion of improved Quality of Life. (Though as will later be shown, Emotions Composite was only a sign of a successful intervention.)

Time to follow up The number of days between text 3 and follow up was significantly and negatively related to outcome measure Days of physical health ($r = -.60$, $p < .01$), indicating that a longer follow up period was associated with reporting fewer physical symptoms. No correlations were found to the other outcome measures.

Psychological health

The mean score of Quality of Life (as measured by BBQ) in the final sample (n = 20) ($M = 62.40; SD = 19.82$) was slightly above the mean score found in a Swedish population ($M = 60.08; SD = 15.72$) (Frykheden, 2014). In order to test any significant differences on condition or within subjects on Quality of Life/BBQ, a Repeated Measures ANOVA was
Table 4
Correlations between baseline Optimism, Pronoun flexibility, Emotions Composite, Time to follow up, and Outcome measures change-results (n = 20)

<table>
<thead>
<tr>
<th></th>
<th>LOT-R pre</th>
<th>SDjag</th>
<th>SDjag t1t2</th>
<th>SDjag t2t3</th>
<th>MeanSDjag t12&amp;t23</th>
<th>Jag/words t3-t1</th>
<th>Emotions Composite</th>
<th>Days btw WM1 &amp; WM2</th>
<th>Days btw Text 3 &amp; follow up</th>
<th>AOSPAN Change</th>
<th>BBQ Change</th>
<th>Physical symp Change</th>
<th>Psych symp Change</th>
<th>Restricted Activities Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOT-R pre</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDjag</td>
<td>-0.06</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDjagt1t2</td>
<td>0.02</td>
<td>0.83  ***</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDjagt2t3</td>
<td>-0.04</td>
<td>0.87  ***</td>
<td>0.68 **</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MeanSDjag t12&amp;t23</td>
<td>-0.02</td>
<td>0.93  ***</td>
<td>0.90 ***</td>
<td>0.93 ***</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jag/words t3-t1</td>
<td>-0.09</td>
<td>0.46  *</td>
<td>0.08</td>
<td>0.26</td>
<td>0.19</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotions Composite</td>
<td>-0.09</td>
<td>-0.31</td>
<td>-0.19</td>
<td>-0.43</td>
<td>-0.36</td>
<td>-0.14</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days between WM1 &amp; WM2</td>
<td>-0.07</td>
<td>0.14</td>
<td>0.21</td>
<td>0.14</td>
<td>0.18</td>
<td>0.20</td>
<td>0.24</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days between Text 3 &amp; WM2</td>
<td>0.18</td>
<td>-0.16</td>
<td>-0.08</td>
<td>-0.15</td>
<td>-0.14</td>
<td>0.12</td>
<td>0.16</td>
<td>0.77 ***</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOSPAN Change</td>
<td>0.36</td>
<td>0.48  *</td>
<td>0.46 *</td>
<td>0.59 **</td>
<td>0.58 **</td>
<td>0.13</td>
<td>-0.06</td>
<td>0.27</td>
<td>0.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBQ Change</td>
<td>-0.22</td>
<td>-0.25</td>
<td>-0.16</td>
<td>-0.39</td>
<td>-0.32</td>
<td>0.05</td>
<td>0.60 **</td>
<td>0.20</td>
<td>0.16</td>
<td>-0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical symptoms Change</td>
<td>-0.00</td>
<td>0.05</td>
<td>0.07</td>
<td>0.12</td>
<td>0.11</td>
<td>-0.24</td>
<td>-0.07</td>
<td>-0.49 *</td>
<td>-0.60 **</td>
<td>0.04</td>
<td>-0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological symptoms Change</td>
<td>-0.20</td>
<td>0.05</td>
<td>0.14</td>
<td>0.07</td>
<td>0.11</td>
<td>-0.11</td>
<td>-0.10</td>
<td>-0.05</td>
<td>-0.18</td>
<td>0.00</td>
<td>-0.10</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restricted Activities Change</td>
<td>-0.02</td>
<td>0.14</td>
<td>-0.05</td>
<td>0.14</td>
<td>0.05</td>
<td>0.04</td>
<td>0.40</td>
<td>0.17</td>
<td>0.01</td>
<td>0.05</td>
<td>-0.08</td>
<td>-0.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05, ** p < .01, *** p < .001
utilized. As seen in Table 5, a significant difference was found on condition ($F = 8.21, p = .01, \eta^2 = .31$). Experiment Group increased their score, from pre-test ($M = 65.30; SD = 16.61$) to post-test ($M = 72.50; SD = 13.99$), while Control Group decreased their score from pre-test ($M = 59.50; SD = 23.13$) to post-test ($M = 55.40; SD = 19.55$). Quality of Life/BBQ was the only outcome measure in the study that gave effect on condition. Writing in the experiment condition increased self-reported Quality of Life.

As mentioned earlier, for all participants in the final sample, correlations were found between BBQ change-results and Emotions Composite; yet, when adding Emotions Composite as a covariate in the Repeated Measures ANOVA, no significant results were found ($p$s > .05) (see Table 6). Emotions Composite thus only confirmed that the intervention was successful.

### Table 5

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of Life (BBQ)</td>
<td>24.03</td>
<td>1</td>
<td>24.03</td>
<td>0.62</td>
<td>0.44</td>
<td>0.02</td>
</tr>
<tr>
<td>Quality of Life (BBQ) * Condition</td>
<td>319.23</td>
<td>1</td>
<td>319.23</td>
<td>8.21</td>
<td>0.01</td>
<td>0.31</td>
</tr>
<tr>
<td>Residual</td>
<td>700.25</td>
<td>18</td>
<td>38.90</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Type III Sum of Squares

### Table 6

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of Life (BBQ)</td>
<td>32.59</td>
<td>1</td>
<td>32.59</td>
<td>0.86</td>
<td>0.37</td>
<td>0.04</td>
</tr>
<tr>
<td>Quality of Life (BBQ) * Condition</td>
<td>3.55</td>
<td>1</td>
<td>3.55</td>
<td>0.09</td>
<td>0.76</td>
<td>0.00</td>
</tr>
<tr>
<td>Quality of Life (BBQ) * Emotions Composite</td>
<td>54.97</td>
<td>1</td>
<td>54.97</td>
<td>1.45</td>
<td>0.24</td>
<td>0.08</td>
</tr>
<tr>
<td>Residual</td>
<td>645.28</td>
<td>17</td>
<td>37.96</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Type III Sum of Squares

### Working memory

The result of WM test AOSPAN at pre-test was subtracted from the result of post-test and a Repeated Measures ANOVA was employed (time x condition) (Yogo & Fujihara, 2008). Significant results were found for AOSPAN Absolute score within subjects ($F = 7.78, p = .01, \eta^2 = .30$), but not on condition (see Table 7). Thus, WM results improved, but not due to the EW task.

**First-person singular pronoun.** When adding SDjagt2t3 (which had the strongest correlation to AOSPAN change-results) as a covariate in the Repeated Measures ANOVA, the effect was even stronger for this covariate ($F = 8.95, p = .01, \eta^2 = .35$) (see Table 8). Thus, the variation in percentage use of “jag” (“I”), especially between text 2 and text 3, influenced the AOSPAN post-test results in both conditions.
Table 7
RM ANOVA: Within Subjects Effects on AOSPAN Absolute score

<table>
<thead>
<tr>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>AOSPAN</td>
<td>220.90</td>
<td>1</td>
<td>220.90</td>
<td>7.78</td>
<td>0.01</td>
</tr>
<tr>
<td>AOSPAN * Condition</td>
<td>0.90</td>
<td>1</td>
<td>0.90</td>
<td>0.03</td>
<td>0.86</td>
</tr>
<tr>
<td>Residual</td>
<td>511.20</td>
<td>18</td>
<td>28.40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Type III Sum of Squares

Table 8
RM ANOVA: Within Subjects Effects on AOSPAN Absolute score with SDjagt2t3 as covariate

<table>
<thead>
<tr>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>AOSPAN</td>
<td>4.60</td>
<td>1</td>
<td>4.60</td>
<td>0.240</td>
<td>0.63</td>
</tr>
<tr>
<td>AOSPAN * Condition</td>
<td>8.04</td>
<td>1</td>
<td>8.04</td>
<td>0.419</td>
<td>0.53</td>
</tr>
<tr>
<td>AOSPAN * SDjagt2t3</td>
<td>171.90</td>
<td>1</td>
<td>171.90</td>
<td>8.953</td>
<td>0.01</td>
</tr>
<tr>
<td>Residual</td>
<td>307.21</td>
<td>16</td>
<td>19.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Type III Sum of Squares

Number of days to AOSPAN post-test. The mean number of days between the AOSPAN pre-test and post-test was 24 days ($SD = 4.37$), and, as mentioned, the mean number of days between the last text, Text 3, and AOSPAN post-test was 19 days ($SD = 4.46$). After dividing the sample through a Median split on AOSPAN change-results, an Independent samples $t$-test was run to test differences between Median split groups on number of days. No significant differences was found (all $ps > .05$). Additionally, some participants took the AOSPAN post-test during Christmas vacation (defined as 2017-12-23 to 2018-01-07), which could somehow have influenced the results, but an Independent samples $t$-test showed no significant results ($p > .05$). Thus, neither the number of days between AOSPAN pre-test and post-test, nor between Text 3 and AOSPAN post-test, nor the impact of Christmas vacation, could explain the AOSPAN change-results.

Reliability of the self-administered working memory test. Since the participants took the WM test AOSPAN on their own, it was considered important to examine if the results seemed reliable. As mentioned earlier, three outliers were identified through Interquartile Range. The three identified outliers had extraordinary change-results: in percentage 560 %, -58 %, and -37 % respectively. (Noteworthy is that the two outliers with the most extreme change-results had been registered for taking the AOSPAN test twice, at either pre-test or post-test, and within a time span of 13 minutes, respectively 4 hours.)

Redick et al. (2012) found the mean on AOSPAN Absolute score in a sample aged 17-35 years ($n = 6236$) to be 42.04 ($SD = 17.67$). The final sample ($n = 20$) in the current study (aged 21-69 years; $M = 45.45$ years; $SD = 16.08$) received a baseline mean of 46.05 ($SD = 18.85$) on AOSPAN Absolute score. Thus, the mean of the current sample was a bit higher, compared to Redick et al. (2012), while the SD was rather similar. Further, in comparison with the homogenous samples in former studies of EW and WM, consisting of college students (Klein & Boals, 2001b; Yogo & Fujihara, 2008; Kellogg et al., 2010), the current sample was heterogeneous, both according to occupation and age span.

In Table 9, results for AOSPAN pre-test, AOSPAN post-test and AOSPAN change-results can be seen. Correlations between the results of AOSPAN pre-test and post-test were run and found highly correlated ($r = .90$, $p < .001$). A rank-order comparison of the participants’ pre-test and post-test results (Redick et al., 2012) was made to see if the AOSPAN change-results were reliable. The absolute difference between the participants’ pre-
test and post-test rank-orders seemed quite stable ($M = 1.30; SD = 1.03$, Variance: 1.06, Range = 4). Finally a Reliability statistics test gave a Cronbach’s $a = .96$. Thus, the self-administered AOSPAN test in the current study appeared to be reliable.

### Table 9
**Descriptive statistics for the final sample on AOSPAN Absolute score**

<table>
<thead>
<tr>
<th></th>
<th>AOSPAN pre</th>
<th>AOSPAN post</th>
<th>AOSPAN Change</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>46.05</td>
<td>50.75</td>
<td>4.70</td>
</tr>
<tr>
<td><strong>Std. Error of Mean</strong></td>
<td>4.22</td>
<td>4.55</td>
<td>1.64</td>
</tr>
<tr>
<td><strong>Median</strong></td>
<td>53.50</td>
<td>58.50</td>
<td>4.00</td>
</tr>
<tr>
<td><strong>Std. Deviation</strong></td>
<td>18.85</td>
<td>20.33</td>
<td>7.34</td>
</tr>
<tr>
<td><strong>Variance</strong></td>
<td>355.4</td>
<td>413.3</td>
<td>53.91</td>
</tr>
<tr>
<td><strong>Skewness</strong></td>
<td>-0.93</td>
<td>-0.90</td>
<td>-0.17</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>-0.08</td>
<td>-0.12</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Minimum</strong></td>
<td>3.00</td>
<td>4.00</td>
<td>-11.00</td>
</tr>
<tr>
<td><strong>Maximum</strong></td>
<td>70.00</td>
<td>75.00</td>
<td>19.00</td>
</tr>
</tbody>
</table>

### A comparison to WM results in former studies of EW and WM

In Table 10 can be seen a comparison to the percentage increases in the earlier studies of EW and WM. The increase of 9% in EW in the current study, was rather similar to earlier studies, but the increase in CW in the present study was larger. While the CW increase of OSPAN results in former studies, varied between 1 to 5% (Klein, 2002, referring to Klein & Boals, 2001b; Yogo and Fujihara, 2008; Kellogg et al., 2010), the increase of AOSPAN results in CW in the present study was 12%.

### Table 10
**Comparisons to the percentage increase of WM results in Klein & Boals (2001), Yogo & Fujihara (2008), and Kellogg et al. (2010), on condition, from baseline to follow up 2**

<table>
<thead>
<tr>
<th></th>
<th>Expressive Writing</th>
<th>Positive Writing</th>
<th>Control Writing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Klein &amp; Boals</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2001b), Exp. 1^</td>
<td>6 %</td>
<td>No group</td>
<td>3 %</td>
</tr>
<tr>
<td><strong>Klein &amp; Boals</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2001b), Exp. 2^</td>
<td>11 %</td>
<td>4 %</td>
<td>3 %</td>
</tr>
<tr>
<td><strong>Yogo &amp; Fujihara</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2008)^</td>
<td>6 %</td>
<td>2 %</td>
<td>5 %</td>
</tr>
<tr>
<td><strong>Kellogg et al.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2010)^</td>
<td>5 %</td>
<td>No group</td>
<td>1 %</td>
</tr>
<tr>
<td><strong>Lindgren</strong>^</td>
<td>9 %</td>
<td>No group</td>
<td>12 %</td>
</tr>
</tbody>
</table>

^ CW instruction from Pennebaker et al. (1990), ^^ CW instruction from Klein & Boals (2001) Exp. 2.

*Note. Results from OSPAN follow up after 7 weeks (Klein & Boals, 2001b), after 5 weeks (Yogo & Fujihara, 2008), after 3 days (Kellogg et al., 2010) and results from AOSPAN after a little more than 2 weeks (Lindgren) - from the last day of the writing interventions in all studies.*

### Self-reported illness

A Repeated Measures ANOVA showed no significant results for change-score for Days of physical symptoms, Days of psychological symptoms, or Days of restricted work
ability or daily activities due to physical or psychological symptoms (all \( p > .05 \)). (As mentioned earlier, significant correlations for the whole final sample (\( n = 20 \)) were found between Time from intervention to follow up and Days of physical symptoms.)

**Comments from the final sample self-reports**

Finally follows observations from the open questions (voluntary comments) in the final sample self-reports. In the experiment condition (\( n = 10 \)) all ten participants wrote comments in Self-report 2, while seven of them commented in Self-report 3. In the control condition (\( n = 10 \)) seven individuals wrote comments both in Self-report 2 and Self-report 3.

*Negative comments from Experiment Group:* In Self-report 2, one participant reported an increased worry/self-criticism after writing, due to difficulties formulating in short time on big questions. In Self-report 3 one participant reported being a bit nervous knowing someone else would read the texts, but that it still was liberating to write. In Self-report 3 one participant reported getting stress reactions from taking the WM test. Three of the participants expressed that the time to write was too short. Two participants wished they had written all three texts about the same subject, to get deeper down. Three participants commented that the writing instructions were quite complicated. Two participants reported difficulties with the different types of e-mails and having technical problems attaching the texts.

*Negative comments from Control Group:* Four participants expressed disappointment in Self-report 2 on being instructed not to write about emotions.

*Positive comments from Experiment condition:* Two participants commented that it was fun and liberating to write. Two participants commented that the writing was good for the soul, respectively for the thoughts and feelings. Two participants commented getting closer to what has happened and getting new insights, respectively letting the inner feelings come out and seeing oneself more clearly. One participant described inspiring others to write, after participating in the experiment.

*Positive comments from Control condition:* One participant described that the writing worked as a way to reflect on what one does in a day, as well as on the stress one feels. Another participant expressed appreciation in describing, planning and visualizing what to do the next day. One participant reported getting aha experiences from the counting of experienced life events, reflecting on how many difficult things the person had experienced. One participant neutrally noted being in the control condition.

**Discussion**

In Expressive Writing (writing emotionally about trauma or stressful events) significant associations between pronoun flexibility and measures of physical health have been documented (Campbell & Pennebaker, 2003). Furthermore significant correlations has been found between trauma writing and improvements in working memory results (e.g., Klein & Boals, 2001b). The main aim of the present study was to combine these two tracks in EW – pronoun flexibility and WM results – to examine whether they were related. Earlier studies in these two EW research tracks have explored the impact of linguistic markers, especially cognitive words, but until now, no study seem to have investigated the impact of pronoun flexibility on WM results.

Somewhat surprisingly the results in the current study was incongruent with earlier studies of EW and WM, since no significant results on outcome measure WM scores (AOSPAN Absolute score) were found on condition. But the final sample as a whole (\( n = 20 \)) significantly improved the results at follow up on WM test, and the effect size was large.
(Cohen, 1988). Adding measures of pronoun flexibility to the analysis of WM test results enhanced the effect size, indicating that the pronoun flexibility was crucial to the results, and that a practice effect could not alone explain the increased AOSPAN results. Pronoun use is a question of perspectives and it was surprising that the change in use of “jag” (in English “I”) was not exclusively found in EW, but also in Time management CW.

Pronoun flexibility was measured by standard deviations of the percentage of “jag” (“I”) from the number of words in the three texts. The strongest correlations were found for the measure of SD of the percentage of “jag” for the second and third text: “SDjagt2t3” and for “MeanSDjagt1t2&t2t3” (i.e., the mean of the SDs for text 1 and 2, and for text 2 and 3, respectively: a measure inspired from an earlier analysis on pronouns with Latent Semantic Analysis; Campbell & Pennebaker, 2003). Since SDjagt2t3 received slightly higher effect sizes and was more easily calculated, this measure was selected. No other significant correlation between flexible use of pronouns and outcome measures was found.

The other outcome measures in the current study was psychological health and self-reported illness. The psychological outcome measure consisted of self-reported Quality of Life (as measured by BBQ), which resulted in the only significant difference on condition. The EW group significantly increased its results, while the CW group decreased its score, and the effect size was large (Cohen, 1988). In the current sample, writing about difficult experiences in the experiment condition improved self-reported Quality of Life at follow up.

No significant results was found on self-reported illness (Days of symptoms and of restricted activities) on conditions. Noteworthy is though, that in the current experiment - with a follow up period of a little over two weeks – there was a significant negative association between Days of physical symptoms and the time between Text 3 and follow up. This indicates that, in the whole final sample, a longer time passing between intervention and follow up was connected to a decrease of physical symptoms. In the study of pronouns and health (Campbell & Pennebaker, 2003), follow up periods were considerably longer; 2-4 months, hence it would have been interesting to follow the self-reported illness for a longer period of time.

Finally the level of baseline optimism did not affect any results within the final sample, incongruent with the results of the work of Cameron and Nicholls (1998).

**The working memory scores.** After the identification and exclusion of three outliers reliability tests within the final sample (n = 20) demonstrated the self-administered AOSPAN test results to be reliable. The final sample increased its mean AOSPAN Absolute score with 4.7 points which is a mean increase of 10 %. For Experiment group the rise was 9 % and for Control group 12 % (which was a non-significant difference). Noteworthy is that while 4.7 points was the mean increase, the change-results varied from -11 points to 19 points. The question is what practical impact the increase of AOSPAN might have in ordinary life, e.g., in studies, in problem solving, in controlling responses? The lack of age-related norms complicates the evaluation of AOSPAN results.

In the current study, there was a bigger mean improvement in WM results, especially concerning the increase in CW condition. While the increases in CW in earlier studies were 1 up to 5 % (Klein, 2002, referring to Klein & Boals, 2001b; Kellogg et al., 2010; Yogo and Fujihara, 2008), the increase in the current study was 12 %. Both Klein and Boals (2001b) and Kellogg et al. (2010) interpreted their increased results in control condition as a practice effect. The generally larger increase on WM results in the current study might partly be explained by four differences between the WM outcome measures used. The former studies used OSPAN, while the current Internet study utilized the automated version AOSPAN, which are correlated to each other with a small to medium effect size. Additionally AOSPAN results are found less stable concerning test-retest reliability, which most probably was
further enhanced in the present study, since the same presentation of letter list was used at both pre-test and post-test. Moreover, the samples in earlier studies were homogenous, consisting exclusively of college students (Klein & Boals, 2001b; Yogo & Fujihara, 2008; Kellogg et al., 2010), while the current sample was heterogeneous, both in occupation and age. A final difference was writing instructions. In the study of Klein and Boals (2001b; Experiment 2), before assigned to condition, all participants briefly defined and objectively recalled the negative event to write about, while in Kellogg et al. (2010) all participants objectively wrote and recalled for 15 minutes. Maybe at least some part of the difference between conditions in these two studies might be due to CW group activating a negative event, without processing it? Kellogg et al. (2010) reported that in CW a rise of intrusive thoughts was related to a rise in avoidance. Klein and Boals (2001b) and Yogo and Fujihara (2008) did not ask participants in advance to recall a negative event. They used an EW instruction about coming to college and a CW instruction of objectively describing a) what they had done yesterday, b) today, and c) a recent social event (Pennebaker et al., 1990).

**Time management control writing.** The most common CW instruction is Time management. But the advantage of EW over Time management CW, is not always gained. Earlier studies have found connections between stress level and performance on WM tests (Klein & Boals, 2001a; Banks & Boals, 2017). Thus, both the experiment condition and the control condition in the present study appeared to have the potential of decreasing stress, but probably in different ways. The EW might have reduced stress levels by the processing of difficult experiences and/or by self-regulation (or self-affirmation or cognitive reappraisal/flexibility). The Time management CW, with its different time perspectives (yesterday, today, tomorrow) (Klein & Boals, 2001b, Experiment 2), might have enhanced reflection on how time has been spent and thus strengthened the motivation to plan for time to come, which usually is beneficial in reducing stress. Still, since the CW instruction was collected from Klein & Boals (2001b), and also used by Kellogg et al. (2010) (personal e-mail communication, 2018-05-14), why did not this hypothesized Time management effect occur also in their experiments? It is 16 years (Klein & Boals, 2001b) and seven years (Kellogg et al., 2010) respectively, between their experiments and the current experiment in 2017. Maybe the levels of stress generally has increased during these years, generating a greater need for Time management? Maybe there are differences in stress levels between USA (where these two former studies were performed) and Sweden? Maybe the current sample was more stressed than the samples of Klein and Boals (2001b) and Kellogg et al. (2010)? Obviously, assessment of self-reported stress - with reliable instruments - is necessary to examine if the Time management CW is associated with lowered levels of stress, and further if those levels of stress are associated with flexible use of the pronoun “jag” (“I”), and improved WM results, respectively?

In a study of older caregivers Mackenzie, Wiprzycka, Hasher, & Goldstein (2007) to their surprise found null-results in the EW group on psychological and physical health outcome measures, but improvements in the CW group. The Time management CW seemed to be beneficial for stressed participants, possibly helping them to plan and solve problems, etc. In an online study of participants with mood disorders, Baikie, Geerligs, and Wilhelm (2012) found that not only the experiment groups (EW and positive writing), but also the control group (Time management writing) improved on measures of psychological and physical health. As one possible explanation, Baikie et al. (2012) referred to the resemblance of Time management with activity scheduling interventions, which are used to treat depression (Mazzucchelli, Kanter, & Martell, 2016). To examine this further, Baikie et al (2012) recommend including a control group who does not write at all, or who writes about something else than time management. The latter was done by Mackenzie et al. (2007), who
concluded that not EW, nor history-writing CW, but only Time management CW resulted in improved health. Examples of other control-writing instructions are to objectively and detailed describe things (Pennebaker & Beall, 1986) or places (Guastella & Dadds, 2006). Baikie et al. (2012), whose sample consisted of participants with mood disorders, proposes that depressed participants might benefit of writing first in Time management CW, then in EW and finally, as they hopefully have improved their well-being, in positive writing.

A last reflection on Time management CW. Maybe writing about what happened yesterday and the present day, elicits more references to other persons, while writing about plans for tomorrow elicits more first-person pronoun-perspective, which might empower and boost the feeling of being an actor, which might in its turn reduce stress levels?

**Expressive Writing instructions.** As mentioned, only one significant result on condition was found in the current study: self-reported Quality of life. According to the Matching hypothesis (Engebretson, Matthews, & Scheier, 1989), a match between interventions and the participant’s natural coping style is required to get the best effect out of an intervention. Maddalena et al. (2014) drew the conclusion that the EW instruction would need to be targeted to the characteristics of the emotional processing of the writer. Furthermore Pennebaker and Chung (2011) recommend writing instructions that allow participants to write about whatever they feel the urge to write about. Probably some participants in the present experiment group would have benefited more if encouraged to write in a manner of their own choice, i.e., those high in emotional processing might have benefited more of writing 20 minutes about their most upsetting experience, without the self-regulating part (which some also wished for in their comments in Self-report 2). But because Cameron and Nicholls (1998) found that pessimists gained only from writing in the self-regulating condition and while at least one study have found that EW about trauma can be contraindicated for people unaccustomed to express their feelings (Niles et al., 2013) - the instructions in the current online study was chosen with the purpose of minimizing the risk of writing being harmful (Tatti, 2007). Likewise, some participants may have gained more from writing in an altogether positive way, but since WM was the main outcome measure, and since improvement of WM is associated only with trauma writing (Klein & Boals, 2001b; Yogo & Fujihara, 2008), a positive writing condition was not a relevant option in the current study.

In summary: the lack of significant differences between conditions on working memory results could partly be due to a non-optimal EW instruction, partly to a surprisingly beneficial Time management CW instruction. It is difficult to determine how big share of the improvement was due to practice effect, but within the current sample it seems clear that pronoun flexibility had impact on the WM results.

**Limitations**

The attrition rate in the current study was as high as 66 %. Probable reasons are: a) that participating in the current study was quite demanding (containing eight parts and requiring at least two hours of engagement during a period of approximately three weeks), b) that the time for the experiment was close to, or even during, Christmas vacation, and c) that no incentives (no payment or course credits) were offered to participants after having fulfilled the experiment. In other Internet studies (where incentives have been paid, sum of money have drawn lots, or at least course credits have been offered), there still have been attrition rates of e.g., 23 % (Tatti, 2006), 35 % (Lee et al., 2016), 65 % (Cutts, 2009), or 74 % (Baikie et al., 2012).

Another reason for attrition was that some participants experienced technical problems (especially difficulties attaching the document to the secure mail when attempting
to e-mail the texts) and therefore did not complete. Sheese et al. (2004) consider the need of technical competence among participants being a limitation in online studies. There were also technical problems with the online AOSPAN test, where some participants did not receive results from either pre- or post-test. Also Baikie et al. (2012) reported (other) technical problems, in their online study, leading to loss of data. Online studies can be vulnerable. An additional limitation was the lack of control of the environment during the WM tests. Although instructions asked the participants to do WM post-test under similar circumstances as WM pre-test, the circumstances might have differed at pre- and post-test, because of disruptions, technical problems, etc.

Due to a radically altered work situation at my workplace, the administrator engaged in this study had a very restricted time to administrate the study, even doing some administration outside working hours. After the change in the work situation the participants were sent the e-mails with the pre-writing parts in the due order, but at the same day. Participants were instructed to do the parts in the study in the stipulated way (see Table 1) - and this worked in all but one case. It would have been advantageous to perform a small pilot study before the current (pilot) study, to be able to check how the mailed instructions worked, and if the technical skills needed for attaching the documents was appropriate, and then adjust the instructions according to identified difficulties, choose another technical solution, etc.

One participant reported having a stress reaction due to the WM test; thus a warning should have been issued in the informed consent about the risk of having a stress reaction. In the current study there was only a single item on stress, while a baseline assessment of perceived stress, with a reliable instrument, could have offered information about correlations between self-reported stress and especially the outcome measures Quality of Life and WM.

Since there were only two test times on outcome measures (baseline and follow up), Individual samples t-test could have been utilized instead of Repeated Measures ANOVA, resulting in a simpler presentation. But nevertheless, with the utilization of Repeated Measures ANOVA the impact of Emotions Composite could be dismissed. Other limitations was that there was only one measure of WM and that the WM test was not normed. Furthermore no analysis of cognitive words were performed, which, if done, could have added interesting information, possible to compare to results of earlier studies.

The participants were not randomly selected, instead recruited through a convenience sample and snowball sampling. Furthermore those who dropped out from the experiment were less optimistic, younger and reported fewer days of physical symptoms, compared to the completers. The final sample \((n = 20)\), when compared to the rest of the participants who filled in Self-report 1, differed only on optimism, with the final sample being more optimistic. Optimism did not affect the results within the final sample, but it obviously impacted on who fulfilled the study. Finally, the final sample was too small for any conclusions to be drawn concerning demographics like age, gender or education. The results from the current study cannot be generalized outside the sample.

**Future research and practical applications**

Replications in larger samples, in Swedish as well as in other languages, are needed to test if the measure of pronoun flexibility presented in the current study, SDjagt2t3 (the standard deviation of the percentage use of first-person singular pronoun “jag” (“I”) in the second and the third text), is significantly associated with improvements in WM outside the current sample. The existing linguistic markers (personal pronouns, cognitive words, emotion words, and verb tense) demand specific software adapted to different languages. If replicated,
the linguistic marker SDjagt2t3 might be an accessible way to measure flexible use of pronouns. Furthermore the use of self-administered AOSPAN needs to be replicated.

There seems to be a potential for the development of EW interventions both within the frames of psychotherapy (Graf et al., 2008), and as a self-help guide (e.g., a waiting list intervention while waiting to see a therapist, or as a way to heal and recover on one’s own). As Frattaroli (2006) states; the act of writing is of low cost, easy done and quite often helpful. An interesting further research would be to first assess baseline and then match the states and traits of the participant with a suitable writing instruction (Maddalena et al., 2014). The baseline assessments (also the outcome measures) could be self-reported degree of: a) depression and anxiety, and/or quality of life and grade of optimism (Cameron & Nicholls, 1998), b) emotional processing (e.g., alexithymia, emotional expressiveness, splitting, repressive coping style) (Lumley, 2004; Baikie, 2008; Maddalena et al., 2014, Niles et al., 2013), c) perceived stress (e.g., to examine whether a self-reported decrease of stress covaries with an increase e.g., in WM capacity and/or Quality of Life), and even d) WM capacity (e.g., Klein & Boals, 2001b), where it would be interesting to also use a normed measure of WM, to be able to better evaluate the importance of the changes in WM results.

After baseline assessments the participants would be suggested one or more of the following writing instructions: 1) writing about a trauma (e.g., Pennebaker & Beall, 1986; Pennebaker et al., 1990), 2) writing about daily stressors (e.g., Tatti, 2006), 3) writing in a self-regulating (Cameron & Nicholls, 1998), or benefit finding (e.g., Guastella & Dadds, 2006; Lu et al., 2012), or self-affirming (Creswell et al., 2007), or reappraising/cognitive processing way (e.g., Guastella & Dadds, 2006; Lu & Stanton, 2010), 4) writing about intensely positive experiences or about one’s best possible self (e.g., King, 2001; King, 2002), or 5) writing from different perspectives (I, you, he/she) (e.g., Anderson & Conley, 2012; Seih, Lin, Huang, Peng, & Huang, 2008). Hopefully this would enhance the benefit of EW.

If the results of the current study (that both EW and time management CW could lead to improved results on WM tests) would be replicated - a practical application could be different learning situations, especially for individuals who find concentrating difficult because of painful memories, worries for the future, or because they are highly stressed. Maybe it would be beneficial to write in both time management CW and EW and positive writing, as suggested by Baikie et al. (2012), or other combinations of writing instructions, depending on baseline assessments?

In conclusion: An interesting development of Expressive Writing would be to, based on baseline assessments, suggest participants suitable writing interventions, but still with the option to choose among different instructions, allowing the choice of a matching coping style. In addition to Time-management Control Writing, it would be informative to add yet another type of CW condition as well as a non-writing condition. If replicated, SDjagt2t3 might function as an easily calculated linguistic marker of pronoun flexibility, to be used in further studies of its relation to outcome measures, especially working memory tests. Replications of using online working memory test AOSPAN is also needed.
References


JASP Team (2018). *JASP (Version 0.8.5)*[Computer software].


Appendix 1

Writing instruction: Expressive Writing

Skrivtillfälle 1

Under tre dagar inom den närmaste veckan (antingen tre dagar i rad, eller utspridda över veckan), skulle jag vilja att du skriver i 20 minuter vid varje tillfälle, om något som känns viktigt och har känslo- mässig betydelse för dig. Det kan vara en negativ livshändelse av något slag, eller något du grubbblar på.

Allt du skriver behandlas fullständigt konfidentiellt.

När du skriver skulle jag vilja att du verkligen utforskar dina innersta känslor och tankar. Vad hände? Vad kände du? Vad tänkte du?

Det du skriver kan handla om dina relationer till andra, som t.ex. föräldrar, kärlekspartners, barn, vänner, släktingar, arbets- och studieamraters, etc.


Du kan skriva om samma ämne alla dagar eller välja olika ämnen.

När du har skrivit i 15 minuter om något som är jobbigt för dig, kan du de sista fem minuterna skriva ner tre saker du kan göra, som kan hjälpa dig att hantera en eller flera av de problem eller utmaningar som du skriver om?

Eller kanske vill du under de sista 5 minuterna skriva något om hur du har tagit dig igenom de jobbiga livshändelsorna? Och hur du kan använda dig av de strategierna framöver också?

Eller kanske vill du skriva om det, trotz allt, har kommit något positivt ur det du har varit med om? Har du t.ex. gjort några förändringar i ditt liv till följd av upplevelserna?

Bekymra dig inte om stavning, satsbyggnad, grammatik, eller liknande. Skriv bara i 20 minuter vid vart och ett av de tre tillfällena.

*  

När en skriver om jobbiga upplevelser är det vanligt att en känner sig tillfälligt låg en stund efteråt, men det brukar gå över inom en timme.

När det gäller Expressivt Skrivande finns en ”Flip Out-Rule: if you feel you will flip out by writing, don’t write.”

Om du känner att något blir för känslo- mässigt jobbigt för dig att skriva om - skriv inte om det. Om du vet att du inte är redo att skriva om ett särskilt smärtsamt ämne - skriv om någonting annat.

Och om det ändå skulle bli för känslo- mässigt jobbigt för dig, medan du skriver: avbryt och ta en promenad eller hör av dig till någon du har förtröende för.

Känn efter vad Du behöver.
Appendix 1, continued

**Skrivtillfälle 2**


De sista fem minuterna kan du skriva om ifall du har provat någon av de tre idéerna du hade vid skrivtillfälle 1, för att lösa problem eller hantera utmaning - och hur det i så fall har gått? Och om du inte har provat; hur kom det sig? Och sedan återigen skriva ner tre saker du kan prova för att lösa problem eller hantera utmaningar? Det kan vara samma idéer eller nya idéer.

Eller kanske vill du under de sista fem minuterna skriva något om hur du har tagit dig igenom de jobbiga livshändelserna? Och hur du kan använda dig av de strategierna framöver också?

Eller kanske vill du skriva om det, *trots allt*, har kommit något positivt ur det du har varit med om? Har du t.ex. gjort några förändringar i ditt liv till följd av upplevelserna?

Bekymra dig inte om stavning, satsbyggnad, grammatik, eller liknande. Skriv bara i 20 minuter vid vart och ett av de tre tillfällena.

**Skrivtillfälle 3**

Samma instruktion som vid skrivtillfälle 2.
Appendix 2

Writing instruction: Control Group

Skrivtillfälle 1

Under tre dagar inom den närmaste veckan (antingen tre dagar i rad, eller utspridda över veckan), skulle jag vilja att du skriver i 20 minuter vid varje tillfälle.

Vid första skrivtillfället skulle jag vilja be dig att i detalj beskriva vad du gjorde igår. Det är viktigt att du beskriver saker precis som de inträffade. Skriv inte om dina egna känslor eller åsikter, utan försök att skriva så objektivt som möjligt.

Bekymra dig inte om stavning, satsbyggnad, grammatik, eller liknande. Skriv bara i 20 minuter vid vart och ett av de tre tillfällena.

Skrivtillfälle 2


Skrivtillfälle 3

Nu skulle jag vilja be dig att i detalj beskriva vad du ska göra imorgon från att du vaknar till att du somnar. Skriv inte om dina egna känslor eller åsikter, utan försök att skriva så objektivt som möjligt.