

March 2017

Evaluating the Implementation and Effectiveness of Reflection Writing

Isabelle Boutet

University of Ottawa, iboutet@uottawa.ca

Marie-Pier Vandette

University of Ottawa, mp.vandette@gmail.com

Sophie-Claire Valiquette-Tessier

University of Ottawa, svali071@uottawa.ca

Follow this and additional works at: https://ir.lib.uwo.ca/cjsotl_rcacea

 Part of the [Curriculum and Instruction Commons](#), and the [Developmental Psychology Commons](#)

<https://doi.org/10.5206/cjsotl-rcacea.2017.1.8>

Recommended Citation

Boutet, I., Vandette, M., & Valiquette-Tessier, S. (2017). Evaluating the Implementation and Effectiveness of Reflection Writing. *The Canadian Journal for the Scholarship of Teaching and Learning*, 8 (1). <https://doi.org/10.5206/cjsotl-rcacea.2017.1.8>

Evaluating the Implementation and Effectiveness of Reflection Writing

Abstract

There is ample theoretical justification for incorporating reflection exercises as a tool for preparing students for life beyond university, yet the utility of such exercises needs to be documented if resources are to be devoted to their implementation. This study describes the implementation and evaluates the effectiveness of a reflection exercise that was introduced in an entry-level undergraduate psychology course. Students completed four periodic reflection journals and submitted an essay summarizing their progress as learner at the end of the semester using examples from the journals to support their reflection. Multiple qualitative analysis methods were used to measure the reflective content of the summary essays. The analyses support the effectiveness of the exercise in promoting reflection on the process of learning including strengths, weaknesses, learning strategies, competence, efforts, and emotions. In addition, reviewing the reflective essays provided the instructor with invaluable insight into the students' experience. We conclude that reflection writing can be incorporated in undergraduate studies to encourage the development of life-long learning skills with reasonable time requirements. Suggestions for modes of implementation as well as future avenues of research are discussed.

Il existe une ample justification théorique à l'incorporation d'exercices de réflexion en tant qu'outils pour préparer les étudiants à la vie après l'université, et pourtant l'utilité de tels exercices a besoin d'être documentée avant de consacrer des ressources à leur mise en oeuvre. Cette étude décrit la mise en oeuvre et évalue l'efficacité d'un exercice de réflexion qui a été présenté dans un cours de premier cycle en psychologie. Les étudiants ont complété quatre journaux de réflexion personnelle au cours du semestre et à la fin du cours, ils ont rédigé un essai dans lequel ils ont résumé leurs progrès en tant qu'apprenants en utilisant des exemples tirés de leurs journaux pour soutenir leur réflexion. Plusieurs méthodes d'analyse qualitative ont été utilisées pour mesurer le contenu réflexif des essais. Les analyses soutiennent l'efficacité de l'exercice pour favoriser la réflexion sur le processus d'apprentissage, y compris les forces, les faiblesses, les stratégies d'apprentissage, la compétence, les efforts et les émotions. De plus, le fait d'avoir examiné les essais de réflexion a fourni à l'instructeur des indications d'une valeur inestimable sur l'expérience des étudiants. Nous concluons que la rédaction réflexive peut être incorporée dans les études de premier cycle pour encourager le développement de compétences d'apprentissage qui serviront tout au long de la vie et ce selon des exigences raisonnables quant au temps requis. Des suggestions de divers modes de mise en oeuvre ainsi que d'avenues pour des recherches futures sont également discutées.

Keywords

journals, reflection, metacognition, evaluation, undergraduate

Cover Page Footnote

This research was supported by a grant from the Center for University Teaching and the Faculty of Social Sciences of the University of Ottawa. We thank Anne-Marie Faltacas for help with data analysis and preparation of an earlier version of the manuscript, Dwayne Schindler and Julie Gosselin for assistance with statistical analysis, Robert Parson for assistance with the development of the reflection exercise, and Cary Kogan for proof-reading the manuscript.

Among a plethora of pedagogical activities available, reflection is often cited as a tool for preparing students for life beyond university (e.g., Black & Williams, 1998; Boud, Keogh, & Walker, 1985; Dewey as cited in Spalding & Wilson, 2002; Falchikov & Boud, 1989; Mezirow, 1990). Indeed, by promoting reflection on learning and attitudes toward learning, professors are hoping to train students to regulate their own learning in an autonomous and reflective manner (Paris & Paris, 2001; Tan, 2008). On theoretical grounds, reflection writing has the potential to positively impact many facets of the learning experience, both cognitively and emotionally (Boud et al., 1985; Mezirow et al., 1990; Wong, Kember, Chung, & Yan, 1995). For example, reflection journals are considered to be a form of metacognition, an ability whose positive impact on learning and academic performance is well documented (e.g., Elwood & Klenowski, 2002; Shepard, 2000; van Kraayenoord & Paris, 1997). Furthermore, reflection journals can enhance awareness of the feelings and attitudes triggered by the learning process and in turn motivate transformation (e.g., Boud et al., 1985; Mezirow, 1990). Finally, reflection journals have the potential to improve student-professor relationships (McDonald & Boud, 2003; Spalding & Wilson, 2002) by offering a safe outlet for personal concerns and frustrations and by providing instructors with invaluable insight into students' experience.

Whether journal writing can actually promote reflection remains controversial. Research on this topic has generally focused on a small number of students registered in professional programs (see review by Mann, Gordon, & MacLeod, 2009 for health professions; Andrusyszyn & Davie, 1997; Duke & Appleton, 2000; Lew, Alwis & Schmidt, 2010; Lew & Schmidt, 2011; McCrindle & Christensen, 1995; Riley-Doucet & Wilson, 1997; Spalding & Wilson, 2002; Stewart & Richardson, 2000; Tyssenaar, 1995; Walser, 2009; Williams, Sundelin, Foster-Seargeant, & Norman, 2000; Wong et al., 1995). Some research has focused on surveying students' and professors' impressions with mixed results (e.g., Maguire, Evans, & Dyans, 2001; Riley-Doucet & Wilson, 1997; Stewart & Richardson, 2000; Walser, 2009). For example, Maguire et al. (2001) surveyed undergraduate geography students' perceptions of the usefulness of reflexive journaling. Students reported completing journals with the purpose of pleasing the professor rather than with a genuine motivation to become better learners. The authors concluded that reflections provided in journals are mainly descriptive and strategic rather than authentic and reflective. Walser (2009) reported contradictory results. Students enrolled in an undergraduate education program were asked to reflect on their performance as well as areas of strength and improvement *via* three short self-assessment exercises in the course of the semester. The exercises consisted of both rating scales and open-ended questions about the material covered in class, factors facilitating their success, as well as insights gained throughout the course of the semester. Students reported that the exercises allowed them to reflect on their progress and performance, motivated them to do well, and gave them the opportunity to provide feedback to the instructor. The instructor also reported a positive experience whereby the exercises helped improve teaching and interactions with the students.

Investigations based on surveys provide a subjective evaluation of the effectiveness and utility of reflection journals. To address this limitation, other researchers have focused on more objective methods to identify the presence and quality of reflections in students' writing (e.g., Lew & Schmidt, 2011; Wong et al., 1995). Most of the research conducted using this approach has focused on students enrolled in professional degrees. For example, Wong et al., (1995) assessed the level of reflection in reflective papers written by nursing students. The papers were analyzed for qualitative content whereby textual elements were categorized according to two pre-established coding systems. Their results indicate that (a) it is possible to categorize the quality of reflection

in journals using content analysis, and (b) the majority of students displayed evidence of meaningful reflection that allowed them to “relate their experience and turn them into new learning opportunities” (p. 54). As far as we know, Lew and Schmidt (2011) have provided the only qualitative investigation of reflection writing in a non-professional context, namely a problem-based learning polytechnic program. This study is also notable because data was drawn from a large number of students ($n = 609$) who had to write daily reflection journals during the course of a semester. The journals were analyzed using qualitative analysis software. Results indicated that the journals contained statements related to critical evaluation of past learning experiences and learning strategies, and also served to summarize what was learned. Students also showed improvements in their reflective skills over the course of the semester. However, correlations between academic achievement and reflective content were either weak or not significantly different from zero. The authors concluded that reflection journals can improve learning in subtle ways that might not be measurable with standard methods of assessing academic achievement.

The goal of the present study was to extend the existing literature on reflection writing by (a) describing the implementation of a reflection journal exercise in an entry-level undergraduate psychology course and (b) investigating with multiple methods of analysis whether writing exercises can promote reflection. The first goal was motivated by the author’s (IB) desire to promote life-long learning among a class of undergraduate students. After consultation with the teaching support centre at her university, reflection emerged as a promising option. Unfortunately, because the practice of reflection is still rare in non-professional undergraduate programs, few references existed to guide implementation. To further complicate matters, many different definitions of reflection and means of implementation exist (e.g., Mann et al., 2009; Tan, 2008). By documenting her experience, the professor wished to provide an additional resource for others who might want to incorporate reflection exercises in their course objectives.

The second goal was to address the inconclusive findings in the literature regarding whether writing actually promotes reflection. Whereas several of the aforementioned studies support the role of journal writing in promoting reflection (e.g., Lew & Schmidt, 2011; Spalding & Wilson, 2002), their generalizability to standard undergraduate curricula is limited. For example, students in the Lew and Schmidt (2011) study were enrolled in a problem-based program that was inherently characterized by self-directed learning and reflection. Furthermore, students completed journals on a daily basis and instructors provided regular feedback. This intervention would be difficult to implement in most undergraduate courses where resources are limited and class sizes are large.

Context

The reflection exercise was implemented with students attending an undergraduate course on child development in the School of Psychology at the University of Ottawa, Canada. The course offered a great opportunity for documenting the role of reflection in a heterogeneous learning environment that draws students from various disciplines and incorporates various pedagogical approaches such as group discussions and debates.

Reflection Exercise

The exercise was designed so it could be incorporated with little additional time commitment. Formal instruction on reflection was not provided. Instead, an organizing framework

(i.e., a list of suggested questions) was provided. This technique can be helpful when introducing novices to the notion of reflection (Spalding & Wilson, 2002). Finally, the reflection exercise was not evaluated for the quality of reflection because of time constraints and to promote honest reflection rather than reflection aimed at providing what would be perceived as desirable to the instructor (Mann et al., 2009; Riley-Doucet & Wilson, 1997; Stewart & Richardson, 2000). To motivate students to participate, the exercise was graded for completion (worth 10% of the final grade).

The reflection exercise consisted of four short periodic journal entries and one final essay. For the journal entries, students were asked to reflect on learning by answering guiding questions (see examples in Appendix A). The questions targeted reflection as defined by Boud et al., (1985), namely, “the intellectual and affective activities in which individuals engage to explore their experience in order to lead to a new understanding and appreciation” (p. 19). As such, the questions were designed to encourage students to reflect on various aspects of their experience as students, including both issues arising from meeting course requirements and participation in class activities. For each journal entry, students had to select one question to be reflected upon in a half-page journal entry to be submitted on-line. Students had to submit two of the journal entries in the first half of the semester and the other two in the second half of the semester, no less than 48 hours apart. Furthermore, to encourage diversity of reflection, students had to choose a different question for each of the journal entries.

At the end of the semester, students wrote a final essay that encouraged them to reflect on their progress as a learner throughout the semester and on future goals to improve their learning. The essay was organized in three sections: (a) where the student started, (b) actions, if any, taken during the semester, and (c) where the student was at now. Students used evidence from their journal entries to support their narrative. In the present study, we analyzed the final essays because reflection is a process that develops over time (Bibauw, 2010; Lew et al., 2010). Furthermore, the questions selected by students for each journal entry were so varied that it would have been difficult to extract consistent themes from these documents.

Methodological Approach

To examine whether the exercise actually promoted reflection, we analyzed the essays provided by two cohorts of students using multiple qualitative and quantitative analysis procedures, a process known as triangulation (Leech & Onwuegbuzie, 2007). Whereas some qualitative analysis procedures like manual coding can be criticized for a lack of objectivity, other techniques that rely on the use of computer software are limited because they ignore the context and make it more difficult to extract meaning from the data (e.g., Leech & Onwuegbuzie, 2008; Smith & Eatough, 2007). To take advantage of the benefits, and mitigate the flaws associated with a particular technique, we analyzed the essays using multiple procedures to increase the representation, objectivity, and reliability of the findings. First, two researchers (MPV, SCV) inductively extracted codes from the essays submitted by a cohort of students enrolled in a section of the course in the winter semester. We prioritized manual coding over automated coding because this type of procedure has the advantage that it permits human interpretation and enhances the ability to extract meaning from the data (Coffey & Atkinson, 1996; Leech & Onwuegbuzie, 2008; Smith & Eatough, 2007). To avoid a biased analysis of the essays, we did not use a pre-established model for coding and the coders were naïve as to the purposes of the analysis. We hypothesized that finding codes related to reflection on learning, and in particular codes corresponding to deeper

levels of reflection, would provide evidence for the effectiveness of the exercise. To evaluate the level of reflection, the codes were classified according to four categories adapted from Boud et al. (1985) and Williams et al. (2000). The four categories were: (a) description of feelings, attitudes, strengths, and weaknesses; (b) analysis of the impact of these descriptions on learning; (c) integrating the analysis to the learning environment (e.g., new understanding or actions towards transformation); and (d) outcomes of the reflection and future behaviors. This allowed us to distinguish strategic reflection from analytical reflection (Lew & Schmidt, 2011; Mann et al., 2009; Spalding & Wilson, 2002). Indeed, if students completed the essays strategically, then one would expect the essays to contain only narrative statements associated with lower levels of reflection (a and b). However, if students learned from the reflective experience, then one would expect the essays to contain analytical statements associated with deeper levels of reflection (c and d).

Second, the essays were imported in an automatic coding software, and the same two researchers identified independently the presence of each code in the essays. This allowed us to calculate code frequencies, which served three goals. First, code frequencies were extracted separately for two separate cohorts to ascertain the external validity of the findings. Second, code frequencies were compared with the students' academic performance. Third, code frequencies were compared to an informal assessment of the essays to examine whether a formal evaluation might be included in future courses. After the course had ended, the instructor assessed each essay with a correction grid linked to the four classification categories listed previously. Students received one point when their reflection included at least one statement related to each of the four category. Based on a subjective evaluation by the instructor, an additional point was given if the essay appeared authentic and honest (1 or 0). Students were unaware that their essays would be marked and the score was not included as part of their overall evaluations for the course.

Third, word count analyses were performed under the assumption that the more frequently a word is used, the more important the word is for the person (Carley, 1993). As such, word count was used to examine whether or not the most frequently used words in the students' narratives were indicative of reflection on the process of learning. Again, word count analyzes were performed separately in the two cohorts to evaluate external validity.

Method

Participants

Students wrote essays as part of the requirements in a course on child development at the University of Ottawa, Canada. The first set of essays was collected from a cohort of 51 students enrolled in the course in a regular 12-week winter semester in 2011 (first cohort). The second set of essays was taken randomly from 25% of students enrolled in the 6-week intensive summer semester in 2012 (second cohort; 20 essays). The sample was a convenience sample based on student enrolment and submission of the reflection exercise. The University's Research and Ethics Board approved secondary use of the students' narratives and grades for the purpose of this study.

Data Analysis

First analysis: Constant comparison analysis. We used a manual constant comparison analysis to inductively extract codes from the essays produced by the first cohort. Two codebooks

were independently extracted from the 51 essays by two researchers (MPV, SCV). The coders were not aware that the exercise was meant to promote reflection on learning when they performed the analysis. The researchers apposed descriptive codes to the sentences in the narratives. After the two researchers each extracted their codebook from the narratives, the principal investigator evaluated the inter-rater agreement of the codes using the following scale: 0 “no agreement”, 1 “little agreement”, 2 “much agreement”, 3 “total agreement” (Denzin, 1978). Only those codes that received either a 2 or 3 rating were kept for the final codebook. During this process, only 5% of the codes were eliminated, indicating good objectivity and reliability for the initial extraction of the codes. Finally, the codebook was organized according to the four levels of reflection listed in the introduction: (a) description, (b) analysis, (c) integration, and (d) outcomes. Two researchers independently assigned each code to one of these four levels. There was generally low agreement regarding which level a given code belonged to and only 50% of the codes were assigned to the same level by both raters. To resolve differences, the authors reviewed several examples for the codes for which there was disagreement and a final classification was agreed upon. Whenever the two raters could not agree, the lower level was always chosen to avoid positively biasing the data to support the hypotheses being tested.

Second analysis: Classical content analysis. NVivo coding software was used to measure code frequencies. The essays were imported in NVivo and each coder independently highlighted segments of the essays that corresponded to one of the codes when they were relevant. The frequency of the segments associated with a code (frequency count) was calculated.

Third analysis: Word count. Word count was performed separately on the data from the two cohorts of students using NVivo software.

Results

Codebook

The codebook can be found in Appendix B. All of the codes that were extracted from the essays were related to learning, emotions and attitudes. Critically, the analysis showed evidence of reflection at all four levels, including five out of thirteen codes relating to the two higher levels which have been described as “deeper” or more “analytical” (Mann et al., 2009). Therefore, the analysis performed by two naïve coders who agreed on 95% of the themes extracted offers evidence that the exercise actually promoted reflection on learning.

Classical Content Analysis

Figure 1 shows the frequency distribution for the sum of the codes for each category and for each participant for both cohorts. Code frequencies for the four categories were consistent across the two cohorts (Table 1). To examine the possible association between code frequency and academic performance, data from all students were divided into two groups: one group of “strong reflectors” where the frequency of all codes combined was above the rounded total average (=8) and another group of “weak reflectors” where the frequency was below this value. The total mark (/100) obtained in the course was then compared across these two groups (Table 2). No significant difference between the groups was observed [$t(69) < 1$].

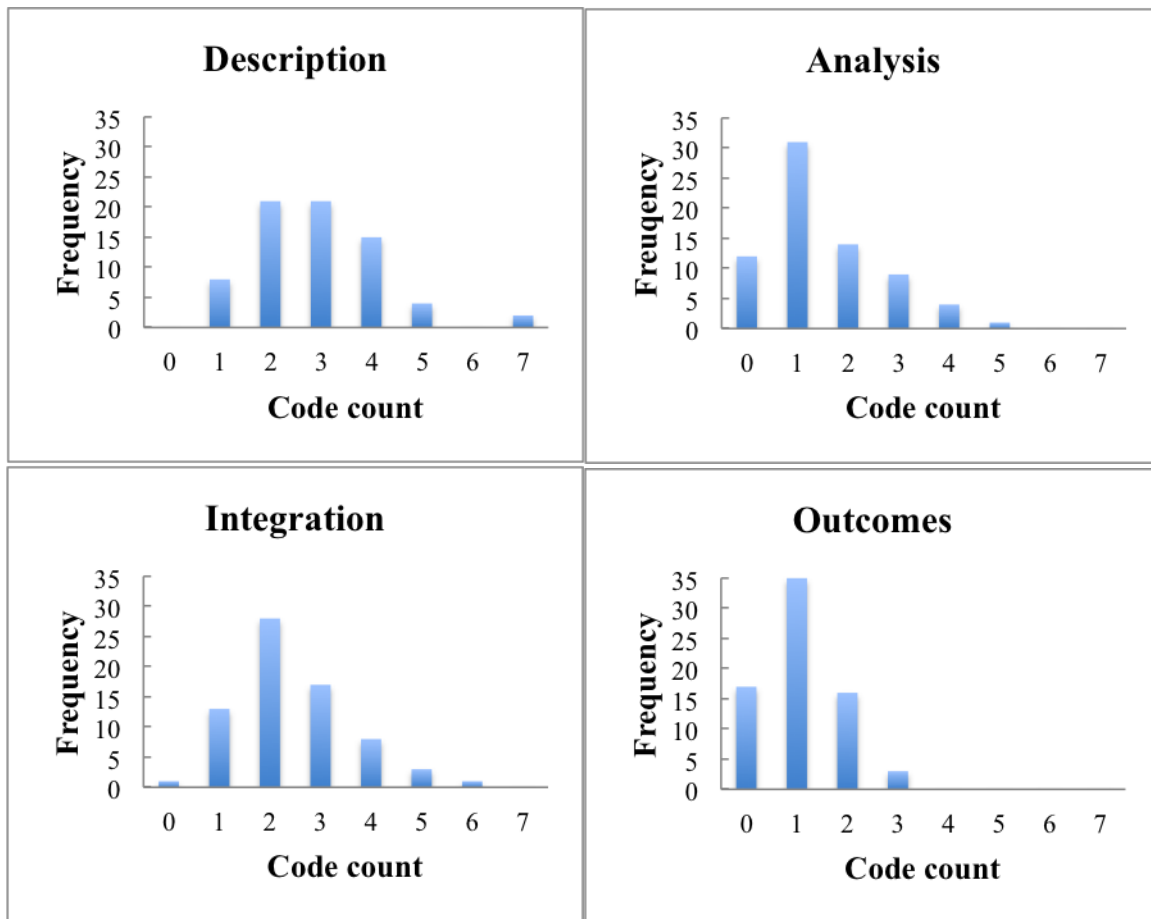


Figure 1. Frequency distribution for the sum of the codes present in each student's essay for each category for both cohorts collapsed.

Table 1
Mean and Standard Deviation (SD) Frequency Count for Each Cohort

	Description	Analysis	Integration	Outcomes	Total
Cohort #1					
Mean	3.07	1.35	2.63	1.18	8.24
SD	1.34	1.07	1.22	0.77	2.13
Cohort #2					
Mean	2.5	1.90	1.95	0.80	7.15
SD	1.05	1.33	0.89	0.83	1.76

Table 2

Students' Standing with Respect to the Mean Frequency of Codes and Academic Performance

Group	N	Average final mark	SD
Strong reflectors	22	76.50	7.86
Weak reflectors	37	75.22	7.89

We also used code frequency data to examine the percentage of the essays that related to the themes extracted in the codebook for each cohort. Table 3 illustrates the percentage of coverage calculated in NVivo. If 100% of the essays' contents were linked to the codes in the codebook, then the coverage would be 100%. Both cohorts showed similar percentage of coverage with the exception of the analysis level where the percentage was higher in the first as compared to the second cohort. These results indicate that a majority of the content of the students' essays were actually related to reflection on the learning experience and that the codes were not exclusively from the lower levels, suggesting that students did not adopt a purely strategic approach when completing the reflections.

Table 3

Mean and Standard Deviation Percent (%) Coverage for Each Cohort

	Description	Analysis	Integration	Outcomes	Total
Cohort #1					
Mean	24	25	22	5	63
SD	10	15	11	7	14
Cohort #2					
Mean	25	12	28	5	82
SD	10	10	17	5	16

Finally, code frequencies were compared to an informal evaluation by the instructor (Figure 2). The goal was to determine whether a subjective evaluation of the essays would be consistent with the more objective evaluation offered by the frequency count. Visual inspection of Figure 2 reveals that there was no consistent relationship between the instructor's evaluation and the overall code frequency counts ($r(70) = 0.11$, $p > .05$). This suggests that the instructor's evaluation did not offer a valid means of assessing the level of reflection in the essays.

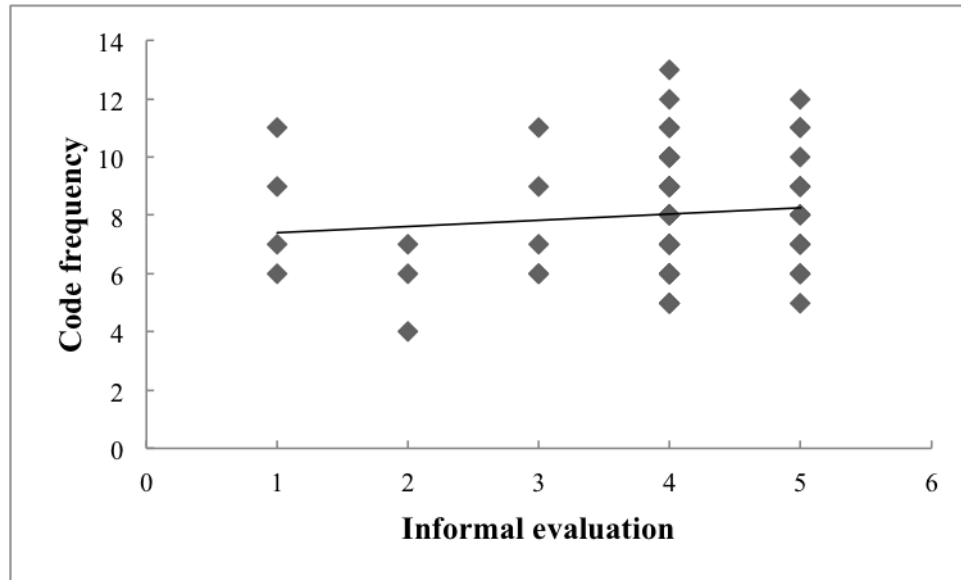


Figure 2. Scatter plot with the instructor’s assessment on the x axis and overall code frequencies on the y axis.

Word Count

The ten most reported words are illustrated separately for each cohort in Figure 3. The most frequent words used were generally consistent across the two cohorts.



Figure 3. Word clouds created on the basis of the word count analysis for the two cohorts separately (left panel: first cohort; right panel: second). The 10 most frequent words are illustrated with the more prominent ones being the most frequent.

Discussion

The goal of the present study was to document and evaluate the implementation of a reflection exercise in an entry-level undergraduate course in psychology. The reflection exercise was developed to promote life-long learning in students in a way that would add little burden to the professor’s workload and therefore provide a feasible pedagogical tool. During the semester, students completed periodic short reflections prompted by a list of questions. At the end of the semester, students completed a one-page summary essay describing their progression during the semester. We analyzed the reflective content of these essay using different methods. We hypothesized that finding codes related to reflection on learning, and in particular codes

corresponding to the deeper levels of reflection, would serve as evidence for the effectiveness of the exercise.

Professor's Experience

By now, the reflection exercise has been included in three additional sections of the same course. The experience suggests that such an exercise could be integrated with reasonable time requirements in undergraduate classes of less than 100 students. The questions used to guide the journal entries are simple and general, allowing its application to a variety of courses and disciplines. With respect to time constraints, approximately 30 minutes per week were devoted to randomly check some of the journals to ensure that the students were following the instructions. Feedback was only provided to students who failed to comply with instructions. At the end of the semester, students submitted their four periodic journals in a printed format with a proof of date completed together with their final essay. Reviewing this material for compliance with the instructions required approximately two to three hours.

One of the most satisfying aspects of incorporating reflection in a course is the feedback it offers, both positive and negative. Reading that students had faced their shyness during group work or developed new organizational skills during the course of the semester gave new meaning to my role as a university professor: I felt a great pride in knowing that I had imparted not only knowledge but also fostered life-long learning.

Was the Exercise Effective in Promoting Reflection?

Some researchers have argued that reflection writing assignments do not promote authentic reflection but instead are descriptive accounts of the students learning activities completed with the purpose of pleasing the professor (Lew & Schimdt, 2006; Maguire et al., 2001). Our analysis suggests that the reflection exercise did promote reflection on learning as defined by Boud et al., (1985), that is “the intellectual and affective activities in which individuals engage to explore their experience in order to lead to a new understanding and appreciation” (p. 19). Indeed, the exercise encouraged students to reflect on their strengths and weaknesses, strategies for improvement, outcomes for future learning, emotions and attitudes. Critically, students did not simply narrate events as would have been expected if they had adopted a strategic approach to the exercise (Lew & Schmidt, 2011; Mann et al., 2009; Spalding & Wilson, 2002). Instead, students produced statements associated with the highest levels of reflection. At these higher levels, the experience of reflection can lead to transformation by encouraging critical examination of events, acquisition of personal knowledge, and the attainment of new perspectives (Wong et al., 1995).

The study also examined whether reflection is associated with academic performance. We divided students from both cohorts as “strong or “poor” reflectors on the basis of their standing with respect to the frequency of codes. No significant difference was observed between the two groups with respect to their final grade in the course. However, our sample size did not provide sufficient power for detecting small or medium effects, making it difficult to draw definitive conclusions on the basis of this result. Other studies have reported similar results with reflection journals have no or little impact on grades (Lew & Schmidt, 2011). Such findings are not surprising because reflection is a process that develops over time and it is unlikely that all students equally matured in this process during the course of the semester. Reflection might not lead to immediate benefits in academic achievement but rather have a more general positive impact on the student

experience by stimulating deep approaches to study, encouraging metacognitive awareness, enhancing lifelong learning, increasing motivation, providing an outlet for the affective aspects of learning, and improving student-teacher relations (e.g., Boud et al., 1985; Elwood & Klenowski, 2002; Kraayenoord & Paris, 1997; Lopez & Kossack, 2007; Mezirow, 1990; McDonald & Boud, 2003; Reynolds & Trehan, 2000; Shepard, 2000; Spalding & Wilson, 2002; Tan, 2008; Williams et al., 2000). It would be interesting for future studies to incorporate valid measures of other aspects of the student experience, for example the Metacognitive Awareness Inventory (MAI; Schraw & Dennison, 1994), to quantify the benefits of reflection.

Some may question whether content analysis is a valid approach for evaluating whether the exercise actually promoted reflection. In a review of past studies on reflection, Mann et al., (2009) suggested that qualitative research approaches are ideal because the study of reflection writing in higher education is still in development. As suggested by Mann et al. (2009), we created the exercise and couched our data analysis within the context of a precise operational definition of reflection as well as pre-established criteria to categorize the depth of reflection. Despite the multiple methodological elements included to ensure that the analysis would offer a valid evaluation of the effectiveness of the exercise, it remains that the present study is limited by its reliance on the students' statements. Indeed, these statements might be influenced by perceptions of safety for revealing personal information or by apprehension at the idea of having the professor evaluate the text. The reality is that qualitative analysis cannot distinguish authentic and honest statements arising from genuine reflection from superficial statements produced to comply with course instructions. Unfortunately, there might not be any methodology that would make this distinction possible. To address this limitation, we analyzed the data on the basis of levels of reflection to distinguish superficial/strategic reflection from analytical reflection and found evidence of the latter.

It is also possible that the reflective skills measured herein were not triggered by the exercise *per se* but simply a manifestation of a competency that some students already possess intuitively (Lew et al., 2010). The distribution frequency graphs (Figure 1) indicate that levels of reflection were normally distributed. More importantly, all students showed some evidence of reflection, suggesting that the exercise can be useful even for those who possess little propensity to reflect in the first place. After analyzing the reflections of students enrolled in a professor education program, Spalding and Wilson (2002) came to a similar conclusion. They remarked that while different students began the semester with diverse proficiency and familiarity with reflective thinking and writing, their intervention was useful for all students by fostering individual patterns of growth. More research is needed to document the potential long-term effects of engaging students in reflection early in their undergraduate degree.

Our results are unlikely to reflect a confirmation bias because the themes were extracted by coders who were naïve to the goals of the study. Still, because manual coding procedures are sometimes criticized for their lack of objectivity, a process of triangulation was used (Leech & Onwuegbuzie, 2007), and automatic analyses were also included to ensure more objectivity. Finally, the results were generalized by comparing two different cohorts. On the basis of our analyses, we conclude that journal writing can promote reflection on the process of learning.

Should Reflection be Evaluated?

The issue of whether or not reflection writing should be graded remains debated (e.g., see discussion by Spalding & Wilson, 2002; Stewart & Richardson, 2000; Williams et al., 2000). The

informal assessment of the essays was not related to the frequency of the codes extracted objectively. Wong et al. (1995) reported experiencing similar problems with Boud's categories, which are comparable to those used here. They suggested instead a classification in line with Mezirow (1990) model whereby students are rated as non-reflectors, reflectors, or critical reflectors on the basis of the presence and type of reflective elements in their writing. The authors stated that this process was "straightforward and reliable" (p. 57). Other means of evaluating reflection writing on the basis of specific criteria have also been proposed (Williams et al., 2000).

While awaiting additional research on the best way to evaluate reflection writing, we suggest that instructors provide only participation/completion marks for submitting the journals and summary essay. Students are more likely to provide superficial reflections aimed at pleasing the professor if they know that they will be graded (Stewart & Richardson, 2000). Moreover, by not grading journals, instructors are more likely to encourage free thought (Williams et al., 2000), responsibility (Riley-Doucet & Wilson, 1997), and insight, the very skills we are trying to promote by incorporating reflection in university. Finally, stepping out of the role of evaluator might also be beneficial for the instructor. Indeed, reading the reflection essays from a perspective of curiosity turned out to be very rewarding because it gave permission to remain open to the students' descriptions of their experience.

Conclusion

Instructors who already have too little time to cover the course material might be reluctant to invest time and effort into incorporating reflection writing in their course objectives without empirical evidence of their utility (Weinstein, Husman, & Kierking, 2000). This study adds to a growing body of research suggesting that writing exercises can promote reflection on learning without formal training or feedback (Lew & Schmidt, 2011; Spalding & Wilson, 2002; Wong et al., 1995). Reflection has the potential to offer an indirect support for success by equipping students with transferable skills such as autonomy, metacognition, critical evaluation, and self-awareness. These benefits, together with the invaluable insight into students' experience that reflection writing offers to instructors, lead us to encourage the use of reflection writing as a means to prepare students for life beyond university.

References

- Andrusyszyn, M. A., & Davie, L. (1997). Facilitating reflection through interactive journal writing in an online graduate course: A qualitative study. *Journal of Distance Education/Revue de l'Enseignement à Distance*, 12(1/2), 103-126.
- Bibauw, S. (2010). Écriture réflexive et réflexion critique dans l'exercice du compte rendu. *Revue Internationale de Pédagogie de l'Enseignement Supérieur*, 26(1), 1-27.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-71. <https://doi.org/10.1080/0969595980050102>
- Boud, D., Keogh, R., & Walker, D. (1985). *Reflection: Turning experience into learning*. London, UK: Kogan Page.
- Carley, K. (1993). Coding choices for textual analysis: A comparison of content analysis and map analysis. In P. Marsden (Ed.), *Sociological methodology* (pp. 75-126). Oxford, UK: Blackwell. <https://doi.org/10.2307/271007>
- Coffey, A., & Atkinson, P. (1996). *Making sense of qualitative data*. Thousand Oaks, CA: Sage.

- Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods* (2nd ed.). New York: McGraw-Hill.
- Duke, S., & Appleton, J. (2000). The use of reflection in a palliative care program: a quantitative study of the development of reflective skills over an academic year. *Journal of Advanced Nursing*, 32(6), 1557-1568. <https://doi.org/10.1046/j.1365-2648.2000.01604.x>
- Elwood, J., & Klenowski, V. (2002) Creating communities of shared practice: The challenges of assessment use in learning and teaching. *Assessment & Evaluation in Higher Education*, 27(3), 243-256. <https://doi.org/10.1080/02602930220138606>
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59(4), 395-430. <https://doi.org/10.3102/00346543059004395>
- Leech, L. N., & Onwuegbuzie, J. A. (2007). An array of qualitative data analysis tools: A call for data analysis triangulation. *School Psychology Quarterly*, 22(4), 557-584. <https://doi.org/10.1037/1045-3830.22.4.557>
- Leech, L. N., & Onwuegbuzie, J. A. (2008). Qualitative data analysis: A compendium of techniques and a framework for selection for school psychology research and beyond. *School Psychology Quarterly*, 23(4), 587-604. <https://doi.org/10.1037/1045-3830.23.4.587>
- Lew, M. D. N., & Schmidt, H. G. (2006). *Reflection upon learning between theory and practice: A focus-group study of tutors' and students' perceptions*. Rotterdam, The Netherlands: Erasmus University.
- Lew, M. D. N., & Schmidt, H. G. (2011a). Writing to learn: Can reflection journals be used to promote self-reflection and learning? *Higher Education Research and Development*, 30(4), 519-532. <https://doi.org/10.1080/07294360.2010.51262>
- Lew, M. D. N., & Schmidt, H. G. (2011b). Self-reflection and academic performance: Is there a relationship? *Advances in Health Science Education*, 16, 529-545. <https://doi.org/10.1007/s10459-011-9298-z>
- Lew, M. D. N., Alwis, W. A. M., & Schmidt, H. G. (2010). Accuracy of students' self-assessment and their beliefs about its utility. *Assessment & Evaluation in Higher Education*, 35(2), 135-156. <https://doi.org/10.1080/02602930802687737>
- Lopez, R., & Kossack, S. (2007). Effects of recurring use of self-assessment in university courses. *The International Journal of Learning*, 14(4), 203-216. <https://doi.org/10.18848/1447-9494/cgp/v14i04/45277>
- Maguire, S., Evans, S. E., & Dyans, L. (2001). Approaches to learning: A study of first year geography undergraduates. *Journal of Geography in Higher Education*, 25(1), 95-107. <https://doi.org/10.1080/03098260125539>
- Mann, K, Gordon, J., & MacLeod, A. (2009). Reflection and reflective practice in health professions education: a systematic review. *Advances in Health Science Education*, 14, 595-621. <https://doi.org/10.1007/s10459-007-9090-2>
- McDonald, B., & Boud, D. (2003). The impact of self-assessment on achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education: Principles, Policy & Practice*, 10(2), 209-220. <https://doi.org/10.1080/0969594032000121289>
- McCrinkle, A. R., & Christensen, C. A. (1995). The impact of learning journals on metacognitive and cognitive processes and learning performance. *Learning and Instruction*, 5, 167-185. [https://doi.org/10.1016/0959-4752\(95\)00010-Z](https://doi.org/10.1016/0959-4752(95)00010-Z)

- Mezirow, J. (1990). How critical reflection triggers transformative learning. In J. Mezirow & Associates (Eds.), *Fostering critical reflection in adulthood* (pp. 1-20). San Francisco: Jossey-Bass.
- Paris, G. S., & Paris, H. A. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36(2), 89-101. https://doi.org/10.1207/S15326985EP3602_4
- Reynolds, M., & Trehan, K. (2000). Assessment: A critical perspective. *Studies in Higher Education*, 25(3), 267-278. <https://doi.org/10.1080/03075070050193406>
- Riley-Doucet, C., & Wilson, S. (1997). A three-step method of self-reflection using reflective journal writing. *Journal of Advanced Nursing*, 25(5), 964-968. <https://doi.org/10.1046/j.1365-2648.1997.1997025964.x>
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460-475. <https://doi.org/10.1006/ceps.1994.1033>
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14. <https://doi.org/10.3102/0013189X029007004>
- Smith, J. A., & Eatough, V. (2007). Interpretative phenomenological analysis. In E. Lyons & A. Coyle (Eds.), *Analyzing qualitative data in psychology*. Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781446207536.d10>
- Spalding, E., & Wilson, A. (2002). Demystifying reflection: A study of pedagogical journal writing. *Teachers College Record*, 104(7), 1393-4121. <https://doi.org/10.1111/1467-9620.00208>
- Stewart, S., & Richardson, B. (2000). Reflection and its place in the curriculum on an undergraduate course: Should it be assessed? *Assessment & Education in Higher Education*, 25(4), 369-380. <https://doi.org/10.1080/713611443>
- Tan, K. H. K. (2008). Qualitatively different ways of experiencing student self-assessment. *Higher Education Research & Development*, 27(1), 15-29. <https://doi.org/10.1080/07294360701658708>
- Tyssenaar, J. (1995). Interactive journals: An educational strategy to promote reflection. *American Journal of Occupational Therapy*, 49(7), 695-702. <https://doi.org/10.5014/ajot.49.7.695>
- van Kraayenoord, C. E., & Paris, S. G. (1997). Australian students' self-appraisal of their work samples and academic progress. *The Elementary School Journal*, 97(5), 523-537. <https://doi.org/10.1086/461879>
- Walser, T. M. (2009). An action research study of student self-assessment in higher education. *Innovative Higher Education*, 34(5), 299-306. <https://doi.org/10.1007/s10755-009-9116-1>
- Weinstein, C. E., Husman, J., & Kierking, D. R. (2000). Self-regulation interventions with a focus on learning strategies. In M. Boekaerts, P. R. Prinrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 72-747). Orlando, FL: Academic Press. <https://doi.org/10.1016/B978-012109890-2/50051-2>
- Williams, R., Sundelin, O., Foster-Seargeant, E., & Norman, G. (2000). Assessing the reliability of grading reflective journal writing. *Journal of Physical Therapy Education*, 14, 23-26.
- Wong, K., Kember, D., Chung, L., & Yan, L. (1995). Assessing the level of student reflection from reflective journals. *Journal of Advanced Nursing*, 22, 48-57. <https://doi.org/10.1046/j.1365-2648.1995.22010048.x>

Appendix A

Examples of questions used to probe reflection in the periodic journals

1. Acquiring Knowledge

Which topic did you have difficulties understanding this week?

- How will you seek clarification on this topic?

2. Contribution to course climate/Participation in group work

Identify in what group setting (small group, pairs, individuals) you work best and why.

3. Practical Skills

Did you manage your time in order to be able to complete the required assignments for this week in a punctual manner?

- If yes, identify strategy(ies) that helped you be punctual
- If no, identify what you could change to improve your punctuality

4. Evaluate your beliefs and attitudes

Was there a topic that made you feel uncomfortable? If so:

- Explore and describe reason(s) why that topic made you feel uncomfortable

5. Identify actions for improvement

Identify one goal you wish to achieve in the upcoming week. Describe how you could achieve this goal

Appendix B

Summary of the twelve codes extracted from the essays (left column) and their definitions and example (right column). The codes are organized according to the four levels of reflection.

(1) Description	
Open to new experiences	<p>Students reflect on their ability to remain open to new experiences and opinions.</p> <p><i>“I had a firm opinion on how things could be done, but have always and will remains always open to suggestions.”</i></p>
Expectations about the course	<p>To expect a particular learning environment.</p> <p><i>“This class was not what I expected it to be. I imagined the class to be strictly lecture and then the exam.”</i></p>
Identification of learning strengths	<p>Students identified their strengths and abilities and how they could use them to succeed.</p> <p><i>“My strengths lie in my ability to choose and synthesize relevant information offered by diverse members regarding the subject at hand.”</i></p>
Identification of obstacles to learning	<p>Students identified obstacles that hindered their learning and performance.</p> <p><i>“The second weakness I had at the beginning of the class was my lack of ability to reflect back what I just learned in class and seeking help right away.”</i></p>
Negative emotion towards the learning environment	<p>At the beginning of the semester, some students were worried or insecure about the group discussions used in the class.</p> <p><i>“Group work has been for me a constant terror since the beginning of my academic career in University.”</i></p>
(2) Analysis	
Awareness of a change in attitude	<p>Change of attitude during the semester.</p> <p><i>“I realized that when I changed my negative attitude into a positive, my grades improved drastically.”</i></p>
Description of strategies used to improve learning	<p>Actions and initiatives the students took to adapt and succeed during the course such as preparing for class, reorganizing a schedule, asking question, etc.</p> <p><i>“When I was writing my second reflection, I discovered that I did not understand what the concept of strabismus was, I sought clarification by re-reading my lecture notes and the PowerPoint presentation.”</i></p>

Metacognitive awareness	<p>Awareness of personal attributes as well as skills that influence learning. Students reported discovering things they ignored about themselves. Some students stated that self-reflection is the best way to become self-aware.</p> <p><i>“I have learned that discussion and group work can be a great way of learning given my personality”</i></p>
(3) Integration	
Identification of new skills acquired during the semester	<p>Students outlined skills they acquired during the semester such as overcoming shyness, communicating more effectively, develop leadership, etc.</p> <p><i>“Having to do groups discussions gave me the confidence needed to succeed in them.”</i></p>
Identification of transferable skills	<p>Students reflected on when and how they could use the skills acquired in this class in different contexts such as other classes, professional environment, etc.</p> <p><i>“I know this skill will open many doors to me in my future aspiring career as a psychologist.”</i></p>
(4) Outcomes	
Satisfaction with progress	<p>Feeling that is reached when a person is happy with improvements, progress, or successful use of a particular learning technique.</p> <p><i>“While I remain unusually anxious, I am satisfied with my progress as I am better able to set aside my emotions and focus on the important topic at hand. “</i></p>
Pride	<p>Many students were proud and empowered by the realization that their accomplishments were the result of their personal efforts.</p> <p><i>“I felt a goal was reached when I found myself able to question the opinions of my peers on important subjects.”</i></p>