

Information technology students' changing perceptions of assessment strategies during pair programming

Jan Hendrik Hahn, Elsa Mentz* and Lukas Meyer

North-West University (Potchefstroom Campus), South Africa

elsa.mentz@nwu.ac.za

Abstract

The aim of the research was to determine to what extent students' perceptions of assessment (peer, self, facilitator and individual) in pair programming situations changed from the beginning to the end of a programming module. In order to achieve this aim, both quantitative and qualitative research methods were used. During the quantitative phase of the research, data was collected by means of a structured questionnaire and analysed statistically. In order to interpret the quantitative findings, qualitative data was collected by means of open-ended items on a questionnaire as well as through interviews. The sample consisted of 20 second year, pre-service student teachers who were majoring in Information Technology at a South African University. Eleven female and nine male students participated, aged 20 to 21 years. The study indicated that students' perceptions of assessment strategies during pair programming situations changed when a variety of strategies were implemented. Before the implementation of pair programming, students indicated that pair assessment methods were not as reliable as individual assessment methods. After having been exposed to different assessment strategies during pair programming, most of the students were in agreement that pair assessment methods could be as reliable as individual assessment methods.

Keywords: Pair programming, peer assessment, self-assessment, individual assessment; facilitator assessment

Introduction and background

Pair programming refers to a programming method where two persons work on the same programming task at the same computer (e.g., Williams & Kessler, 2001; Williams, Wiebe, Yang, Ferzli & Miller, 2002). One person fulfils the driver's role, while the other plays the navigator's role – each role having its own responsibilities (e.g., Hanks, McDowell, Draper & Krnjajic, 2004; McDowell, Hanks & Werner, 2003).

Pair programming is an established practice in the computer programming industry (Conrad, 2000). Therefore, to provide adequately trained staff for this industry, students at tertiary education institutions need to be trained to program in pairs. Williams et al. (2002) as well as Mentz, Van der Walt and Goosen (2008) used pair programming successfully for training computer science students. Evidence of the usefulness of pair programming as a teaching-learning strategy can also be found in the research of Williams and Kessler (2001), Tomayko (2002) and Naggapan, Williams, Ferzli, Wiebe, Yang, Miller & Balik (2003). However, the assessment of students who have been taught to program using a pair programming strategy is generally done individually (Zhao & Jiang, 2009). The possibility that a student may not participate fully during pair programming is one of Urness's (2009) main concerns. He mentions that it is difficult to assess students' skills and abilities while working in pairs. Some students may receive undeserved credit for pair assignments (Katira, Williams, Wiebe, Miller, Balik &

Gehringer, 2004; McDowell et al., 2003). Cliburn (2003) conducted research on students' views of peer assessment and found that, although the majority of participants said that peer assessment made them accountable to their partners, a significant number of them (more than 30%) pointed out that it did not. Hahn (2008) also discovered that some students felt negatively about group assessment.

This raised two questions: How can students' negative perceptions about assessment in pairs be changed? How could concerns about the undeserved credit be addressed? In the light of the aforementioned, the aim of this investigation was to determine to what extent the implementation of different assessment strategies during pair programming changed students' perceptions about assessment in pairs.

Theoretical and conceptual framework

As mentioned in the introduction, it can be a rather complex task to reliably assess the quality of each pair member's contribution during a pair programming situation. In order to address this challenge, some researchers (Elliot & Higgins, 2005; Hanrahan & Isaacs, 2001; Lejk & Wyvill, 2001) included peer and self-assessment strategies.

Peer assessment is done when students, working on the same task, assess each other's contribution to the task. The advantage of peer assessment is that students are involved in the assessment process; by assessing one another they can also assist one another (e.g., Cheng & Warren, 2000, Visram & Joy, 2003). Bloxham and West (2004) reported that students were positive about peer assessment because they could learn from each others' mistakes and identify where improvement was needed. Peer assessment also provided the opportunity for learners to compare their own work with that of others in the group.

Self-assessment, on the other hand, takes place when students assess their own work (Marneweck & Rouhani, 2002) and are encouraged to take responsibility for assessing their own work (Earl, 2003). Although peer and self-assessment may offer some solutions to the problem of unreliable assessment, Cliburn (2003) found that students were still reluctant to award poor marks to group members who did not contribute significantly towards the group task. He also reported that students were even unwilling to give their partners a bad evaluation when they had done nothing during pair programming. Elliot and Higgins (2005) found that although students indicated that peer and self-assessment were effective strategies in assuring fairness and equity in the grading of group projects, peer and self-assessment still held a possibility of unfairness. For instance, some students were reluctant to downgrade individuals who experienced personal problems. This type of subjectivity would be unfair to those students who opted not to share their problems with fellow students.

Hahn (2008) suggests that the facilitator should also assess the work of the pair. *Facilitator assessment* implies that the process and product (quantity and accuracy) of the pair programming task is assessed by the facilitator. It should take place during or after completion of a programming assignment (e.g., Earl, 2003; Hahn, Mentz & Meyer, 2009). However, if the facilitator only assesses the process and product of the pair, no recognition for individual effort within the pair can be given.

Mentz et al. (2008) found that pair programming could be made more effective through the incorporation of principles associated with cooperative learning. One of the principles of this approach to learning is individual accountability (Johnson & Johnson, 2009). In order to foster such accountability within a pair, each member may be required to individually write a program similar to that written by the pair (Mentz et al., 2008). According to Johnson and Johnson (2009), a student's personal accountability can be enhanced if the results of each member's individual assessment were shared by the pair to reflect upon. It seems therefore important that peer and self-assessment in pair programming situations should be followed up by *individual assessment*. Such individual assessments are conducted by the facilitator, just like the pair's work. Knight (2004) holds the view that individual assessment allows students and facilitators to determine individual students' programming efforts and abilities, including their ability to work on their own. In pair programming, individual assessments should follow a pair programming assignment to determine whether both members have individually reached the set outcomes (Mentz et al., 2008).

Researchers such as Marneweck and Rouhani (2002), and Orsmond, Merry and Callaghan (2004) suggest using specific assessment rubrics with clearly formulated assessment criteria for peer, facilitator and self-assessment. The use of rubrics will assist students to focus on the outcomes of an assessment activity (Airasian, 2005).

Drawing on all these studies, Hahn et al. (2009) combined the facilitator assessment of the pair with individual, peer and self-assessment and found that students' individual assessment marks correlated with their peer and self-assessment marks. In the light of the aforementioned, facilitators should not exclude themselves from the assessment process. Students may be under the impression that they have achieved the set outcomes while in actual fact they have not. The role of the facilitator is to give continuous constructive feedback to assist the students in identifying their mistakes and improving their programming abilities (Hahn, 2008). During the feedback process, the facilitator should provide examples of correct solutions, with which students can compare their own solutions (Lambert & Lines, 2000).

This study explores student perceptions. Rathus (2005) defines perception as "an active process in which sensations are organised and interpreted to form an inner representation of the world". It is "a set of mental operations that organizes sensory impulses into meaningful patterns" (Wade & Tavis, 2011:179). The individual constructs or builds on the perceived stimuli (Stenberg, 2006) and is usually sure that what was perceived must be true (Wade & Tavis, 2011). Previous experiences often affect how we perceive the world. Students who had bad experiences with pair assessment in the past may form the perception that all pair assessment is unreliable. To change the perception of people, the stimulus must be changed (Wade & Tavis, 2011). Thus to change perceptions, students must realise that previous experiences with group assessment were not necessarily generalisable, and a new environment for group assessment needs to be created in which the negative stimulus will not be repeated.

The intervention reported in this article attempted to apply self-, peer-, facilitator- and individual-assessment, as discussed above, to pair programming. This investigation sets out to determine to what extent the negative perceptions of students on the reliability of pair assessment could be changed into a more positive direction.

Research design

In order to determine whether students' perceptions of assessment in pair programming situations changed in any way from the commencement to the end of a programming module, both quantitative and qualitative research methods were employed (Creswell, 2009).

The pair programming assessment intervention

Before a pair programming task commenced, participants were informed about the nature of pair programming, with specific emphasis on the roles and responsibilities of both the driver and the navigator. The various assessment strategies were also explained to the students.

The students were expected to complete a programming assignment in pairs each week. The programming tasks were executed in a laboratory setting under the supervision of the facilitator to ensure that all the participants adhered to the guiding principles of pair programming. The participants were allowed sufficient time to complete all programming tasks. Prior to each pair programming task the specific cognitive outcomes that students were expected to achieve were identified and shared. After completion of the pair programming task, the participants completed a peer and self-assessment using a rubric developed according to the specific cognitive outcomes. The facilitator then assessed the quality of the pair programming product using the same rubric. Thereafter an individual test was written on the cognitive outcomes specified. The outcomes of all these assessment strategies were then given back to the members to reflect on individually, as well as in pairs. These strategies were implemented on a weekly basis.

Sampling

Use was made of a convenience sample consisting of the 20 second-year, pre-service student teachers who were majoring in Information Technology education at a South African university. Eleven female and nine male students participated. Their ages ranged from 20 to 21 years. Due to students' absences only 18 of the 20 students completed the structured questionnaire at the beginning and the end of the module.

Data collection

During the quantitative phase of the research, data were collected by means of a structured questionnaire. Qualitative data were collected by means of open-ended items on the same questionnaire, as well as through group interviews.

Instruments

All the participants completed the structured questionnaire at the beginning as well as at the end of the module regarding their perceptions of assessments in pairs. Five weeks elapsed between the beginning and end of the module.

The structured questionnaire consisted of nine items with statements regarding peer, individual, facilitator and self-assessment. The participants had to respond to these statements on a four-point Likert scale. The scale ranged from 1 (no agreement) to 4 (absolute agreement). After each item, a space was provided where participants could qualitatively substantiate their response to the particular item.

In order to gain a deeper understanding of the participants' perceptions of different pair assessment strategies, three group interviews with groups of six participants each were conducted during the course of the module. These interviews were conducted at the beginning, the middle and the end of the module. Different students participated in the three group interviews to avoid a possible contamination effect due to previous participation. Interviews lasted approximately 15 minutes and were conducted in a comfortable environment after class sessions. The four questions posed during the interviews were as follows:

What are your perceptions of self-assessment?

What are your perceptions of peer assessment?

What are your perceptions of facilitator assessment?

What are your perceptions of individual assessment?

Data analysis

For the quantitative Likert scale responses, mean scores were calculated for the participants' responses to each item obtained from the pre- and post-intervention questionnaires. Cohen's effect sizes (Creswell, 2009) were calculated to determine whether practically significant differences existed between the mean scores of the pre- and post-intervention questionnaire data. Interpretation of effect sizes were done according to the following guidelines: 0.2 indicated a small effect, 0.5 a medium effect and 0.8 a large effect (Steyn, 1999).

The substantiations proffered by participants regarding their responses to the individual items in the structured questionnaire and their verbal contributions during the group interviews were analysed qualitatively. A list was compiled of the reasons given by the participants for their responses to the individual items in the questionnaires, and a verbatim transcription was made of the participants' contributions during the group interviews. Thereafter, data segments relating to each other were categorised. These emerging categories were grouped together to form meaningful themes. The services of an independent co-analyst were employed to validate the data analysis process.

Ethical procedures

Permission to conduct the research was granted by the University's ethical committee and students were informed about the proposed research. Students were assured that participation was voluntary and they could withdraw at any stage. They were assured of the confidentiality of the results.

Findings

Table 1 illustrates the change in students' views on pair assessment. These quantitative findings are based on the structured questionnaire responses before and after the experience of the programming module.

Table 1: Students' perceptions of assessment in pair programming situations

Type of assessment	Item no	Item description	Average		Standard deviation		Effect size (d)	Practical significance
			Before \bar{x} 1	After \bar{x} 2	Before	After		
Self assessment	1	It is important that I assess myself to determine whether I have contributed significantly to the development of the program	3.00	3.50	0.91	0.62	0.55	Medium
Peer assessment	2	It is important that the other member of the pair assesses my contribution to the program we have developed	3.22	3.67	0.81	0.49	0.56	Medium
	9	I want to assess my partner to determine his/her contribution to the program we have developed	2.50	3.67	1.20	0.49	0.98	Large
Facilitator assessment	6	It is fair that the facilitator allocates the same mark for a pair programming task to each pair member.	2.61	3.00	1.04	1.03	0.38	Small
	4	Only the facilitator must assess the programming tasks we have done in pairs.	2.39	2.50	0.98	1.25	0.09	Small

Individual assessment	3	I prefer that only the facilitator should assess my individual programming tasks	3.22	3.00	1.06	1.19	-0.18	Small
	5	Facilitator feedback on the programming tasks I have completed is important.	3.61	3.72	0.85	0.57	0.13	Small
Overall: pair assessment	7	Good students are disadvantaged by group assessment methods.	2.44	1.83	1.20	1.10	-0.51	Medium
	8	Group assessment methods are as reliable as individual assessment methods.	2.22	3.06	1.00	0.87	0.84	Large

Table 1 shows that for the overall views of assessment of paired work (items 7 and 8) differences of medium and large practical significance were obtained between the participants’ pre- and post-module perceptions. These perceptions have shifted in favour of pair assessment and away from a decidedly critical stand before the intervention. Table 1 also shows that differences of medium practical significance were obtained for perceptions on self assessment (item 1), and differences of medium and high practical significance for perceptions on peer assessment (items 2 and 9). As far as the perceptions on facilitator and individual assessment are concerned, small differences of practical significance were obtained between the pre- and post-module administration of the questionnaire.

The reasons for students’ perceptions were expressed as motivations for their written responses, and in the verbal contributions in the group interviews. Similar themes emerged from both data sources, particularly perceptions of strengths of self-assessment for error identification; peer assessment for accountability, and facilitator assessment for objectivity. More detailed evidence is provided below.

a. Self assessment

Most of the participants agreed at the beginning (Table 1, item 1, $\bar{x}_1 = 3.00$) and the end (Table 1, item 1, $\bar{x}_2 = 3.50$) of the module that self-assessment was important to determine the

significance of one's own contribution towards program development, because it enhanced *error detection* and helped to *monitor one's own progress*, as expressed by several students below:

*It [self assessment] will help you to gain a good idea about how much work you know;
Otherwise I will not know how much knowledge I have gained;
How else will you know if you can really do the work?*

The interviews conducted during the course of the module revealed a similar finding. Participants explained that self-assessment was very important for purposes of *error identification*. Specific mention was made of the advantages of self-assessment such as identification of one's own errors and limitations, as quotes from several students illustrate below:

*You must learn from your mistakes and detect your mistakes first hand;
You can determine your mistakes;
Now you know your limitations.*

b. Peer assessment

Participants demonstrated some degree of reluctance to assess their partner at the beginning of the module (Table 1, item 9, $\bar{x}_1 = 2.50$). This can firstly be ascribed to their concerns about their ability to assess their partners in a *fair* way and secondly, they were uncertain as to whether they would be able to determine their partners' level of *knowledge acquisition*, as evidenced below.

*I am not good enough to assess the knowledge of my partner;
I will leave it to my lecturer to assess,
I do not understand the work myself.*

At the end of the module participants demonstrated a greater willingness to assess their partners (Table 1, item 9, $\bar{x}_2 = 3.67$). The large effect size ($d = 0.98$) indicates that the participants' willingness to assess their partners' contributions during pair programming had changed significantly during the course of this module, as follows:

If my partner did not do her bit I did not give her any credit; I can now link contribution with knowledge acquisition.

Thus, they felt more comfortable about their ability to assess *fairly* and to determine the *knowledge acquisition* of their partners.

Most of the participants agreed at the beginning (Table 1, item 2, $\bar{x}_1 = 3.22$) and the end (Table 1, item 2, $\bar{x}_2 = 3.67$) of the module that it was important that the other member of the pair should assess them in order to determine their contribution towards program development, because it will reflect on the person's *accountability* as a member.

*...so that I can determine my own contribution;
Then both of us know that we have done our bit;
Sometimes I may think that I am doing enough but my partner thinks differently;*

Everybody contributes equally and not one person does all the work.

Accountability was also a theme that emerged from the interviews. Participants felt strongly that pair members who did not contribute significantly during pair programming tasks would easily be identified when assessed by a peer and as a result peer assessments normally yielded reliable results, as expressed below.

*Then we both know we have done our bit;
Both do the same amount of work. Not only one does all the work.*

This is a very important finding, because one of the points of criticism voiced against pair assessment is that it often results in unreliable results, i.e. credit is given where it is not deserved (McDowell et al., 2003).

c. Facilitator and individual assessment

The results from items 3, 4, 5 and 6 yielded no practically significant differences in terms of the participants' pre- and post-module perceptions on aspects related to facilitator and individual assessment. Although no differences were detected, the pre- and post-module results indicated that participants regarded facilitator assessment as an important part of the assessment process. The *objectivity* theme that emerged from the interviews also lends support to the importance of facilitator assessment.

*The facilitator must assess, otherwise I will give myself too many marks;
The facilitator is objective;
If the facilitator does not assess, then we will give ourselves more marks than we deserve.*

d. Overall pair assessment

At the beginning of the module most participants disagreed or partially agreed that pair assessment methods were as reliable as individual assessment methods (Table 1, item 8, $\bar{x}_1 = 2.22$). This result can be related to participants' concerns about whether pair members' contributions to the program can/will be assessed *honestly*, as illustrated below:

*It is not possible for the facilitator to determine individual group members' contributions; Group members are lenient in their mark allocations because they expect the same lenience in return;
How does a person really know what effort a group member has taken towards solving the problem?*

After completion of the module most participants were in agreement that pair assessment methods were as reliable as individual assessment methods (Table 1, item 8, $\bar{x}_2 = 3.06$). The effect size ($d = 0.84$) indicates that the participants' perceptions of the reliability of group assessment methods had changed significantly during the course of this module. The following remark by one of the participants is apposite:

Now you can determine how much effort your partner has put in - my partners were always honest.

The participants' changed perceptions with regard to the reliability of group assessment methods can be ascribed to their changed views on the *honesty* of group assessments.

At the beginning of the module the participants partially agreed that good students might be disadvantaged by group assessment methods (Table 1, item 7, $\bar{x}_1 = 2.44$). However, at the end of the module most participants took the opposite view (Table 1, item 7, $\bar{x}_2 = 1.83$), justifying their stand as below:

*I think everything was assessed fairly;
Comprehension improved whilst explaining difficult concepts to a partner.*

This finding can possibly be ascribed to the fact that participants' perceptions were initially built upon previous experience of group work where good students were unjustly graded. Their change in perceptions could be as a result of the implementation of the specific group assessment strategies.

Discussion

Different assessment strategies were combined and applied in pair programming. The outcomes of the self, peer, facilitator and individual assessment were communicated to each individual participant and to the respective pairs to which they belonged. The participants therefore understood that they could not allocate unreliable and unrealistic marks to themselves and their partners. The individual test mark provided another good indication of the ability of the students; the participants realised that they could not allocate marks which were too high for the peer or self-assessment, as the marks had to correlate with the outcomes of individual and facilitator assessment. This may explain why students indicated that their perceptions of assessment in pair programming situations changed for the good when self, peer, facilitator and individual -assessment strategies were implemented. When initially asked about this, they were not positive about group assessment methods and complained about possible unfair mark allocation and undeserved credit for members of the group (based on previous experience). When exposed to the assessment strategies described above, most of the students were in agreement that the strategies could yield a reliable indicator of student's individual abilities. They therefore did not complain about unfair mark allocation or undeserved credit for members of the pair after completion of the module.

This finding contradicts other research findings (Elliot & Higgins, 2005) where it is claimed that many students do not wish to be assessed in pairs because they feel that some students may receive undue credit. A possible reason for this difference in perception in this investigation can be ascribed to the fact that peer assessment was only a part of the global assessment strategy; facilitator, individual and self-assessment also contributed to the mark allocation of each student. Students realized that their peers had to be held accountable for their contributions to pair programming tasks and that pair members had to be penalised if they did not contribute to the expected levels and standards. The students understood that different assessment methods were

an important part of pair programming and that they contributed positively to the fairness of the assessment process.

An interesting finding was that students realized, at the end of the module, that the assessment methods used in this research did not disadvantage good students. This finding also stands in contrast to other research findings (Elliot & Higgins, 2005) according to which good students often feel that group work and group assessment methods have a negative effect on their progress and marks. This finding, based on the present investigation, could be ascribed to the fact that individual assessment played an important role in the pair assessment. Students who participated in this study declared that their learning was enhanced by explaining difficult concepts and problems to their partners, something that the pair need in order to achieve a good peer assessment mark. The pair assessment methods described above should therefore be considered for implementation with a view to improving the marks of students in Information Technology.

Conclusion

Students will perceive assessment in pair programming situations in a more positive light if multiple assessment strategies are implemented. These multiple assessment strategies imply that self, peer, facilitator and individual assessment must occur on a regular and continuous basis during pair programming.

Pair members will feel more assured that assessment results are fair when different assessors (self, peer, facilitator) assess their progress on multiple occasions. The role that the facilitator plays in assessing individuals, the pair and the contributions of individuals in the pair must however be emphasized during the assessment process.

Acknowledgement

This research is based on work financially supported by the National Research Foundation (NRF) in South Africa. Any opinion, findings and conclusions or recommendations expressed in this material are those of the authors; the NRF does not accept any liability in this regard.

References

- Airasian, P.W. (2005). *Classroom assessment: concepts and applications* (5th ed.). New York: McGraw-Hill.
- Bloxham, S., & West, A. (2004). Understanding the rules of the game: marking peer assessment as a medium for developing students' conceptions of assessment. *Assessment & Evaluation in Higher Education*, 29(2), 721-733.
- Cheng, W., & Warren, M. (2000). Making a difference: using peers to assess individual students' contributions to a group project. *Teaching in Higher Education*, 5(2), 243-255.
- Cliburn, D.C. (2003). Experiences with pair programming at a small college. *Journal of Computing Sciences in Colleges*, 19(1), 20-29.
- Conrad, B. (2000). Taking programming to the extreme edge. *Infoworld*, 22(30), 61.
- Creswell, J.W. (2009). *Educational research: planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River,: Pearson Education.

- Earl, L.M. (2003). *Assessment as learning: using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin.
- Elliot, N., & Higgins, A. (2005). A self and peer assessment – does it make a difference to student group work? *Nurse Education in Practice*, 5, 40-48.
- Hahn J.H. (2008). Paarassessering teenoor individuele assessering in rekenaarprogrammering. Unpublished Masters dissertation, North West University, South Africa.
- Hahn, J.H., Mentz, E., & Meyer, L. (2009). Assessment strategies for pair programming. *Journal of Information Technology Education*, 8, 273-284.
- Hanks, B., McDowell, C., Draper, D., & Krnjajic, M. (2004). Program quality with pair programming in CS1, *ITiCSE*, 176-179.
- Hanrahan, S.J., & Isaacs, G. (2001). Assessing self- and peer-assessment: the students' views. *Higher Education Research & Development*, 20(1), 53-70.
- Johnson, D.W., & Johnson, F.P. (2009). *Joining together: group theory and group skills* (10th ed.). Upper Saddle River: Pearson Education.
- Katira, N., Williams, L., Wiebe, E., Miller, C., Balik, S., & Gehringer, E. (2004). On understanding compatibility of student pair programmers, *AGSCE*, 7-11.
- Knight, J. (2004). Comparison of student perception and performance in individual and group assessment in practical classes. *Journal of Geography in Higher Education*, 28(1), 63-81.
- Lambert, D., & Lines, D. (2000). *Understanding assessment: purposes, perceptions, practice*. London: Routledge Falmer.
- Lejk, M., & Wyvill, M. (2001). The effects of the inclusion of self-assessment with peer assessment of contributions to a group project: a quantitative study of secret and agreed assessments. *Assessment & Evaluation in Higher Education*, 26(6), 551-561.
- Marneweck, L., & Rouhani, S. (2002). Continuous assessment. In: M. Jacobs, N. Gawe & N.Vakalisa (Eds.), *Teaching-learning dynamics: a participative approach for OBE* (pp. 278-327). Johannesburg: Heinemann.
- McDowell, C., Hanks, B., & Werner, L. (2003). Experimenting with pair programming in the classroom. *ITiCSE*, 60-64.
- Mentz, E., Van der Walt J.L., & Goosen, L. (2008). The effect of incorporating cooperative learning principles in pair programming for student teachers. *Computer Science Education*, 18(4), 247-260.
- Naggappan, N., Williams, L., Ferzli, M., Wiebe, E., Yang, K., Miller, C., & Balik, S. (2003). Improving the CS1 experience with pair programming, *SIGCSE*, 359-362.
- Orsmond, P., Merry, S., & Callaghan, A. (2004). Implementation of a formative assessment model incorporating peer and self-assessment. *Innovations in Education and Teaching International*, 41(3), 273-290.
- Tomayko, J.E. (2002). A comparison of pair programming to inspection for software defect reduction. *Computer Science Education*, 12(3), 213-222.
- Urness, T. (2009). Assessment using peer evaluations, random pair assignment, and collaborative programming in CS1. *Journal of Computing Sciences in Colleges*, 25(1), 87-93.
- Rathus, S.A. (2005). *Psychology concepts & connections*. Canada: Thomson Learning.

- Stenberg, R.J. (2006). *Cognitive psychology (4th ed.)*. USA: Thomson.
- Steyn, H.S. (1999). *Praktiese beduidendheid: Die gebruik van effekgroottes* [Practical significance: The use of effect sizes]. Potchefstroom: PU vir CHO.
- Visram, Z., & Joy, M. (2003). Group assessment for computer science projects. *Proceedings of 4th Annual LTSN-ICS Conference*, 49-53. Retrieved from <http://eprints.dcs.warwick.ac.uk/316/>
- Wade, C., & Tavis, C. (2011). *Psychology (10th ed.)*. Upper Saddle River: Pearson Education.
- Williams, L., & Kessler, R. (2001). Experiments with industry's "pair-programming" model in the computer science classroom. *Computer Science in Education*, 11(1), 7-20.
- Williams, L., Wiebe, E., Yang, K., Ferzli, M., & Miller, C. (2002). In support of pair programming in the introductory Computer Science course. *Computer Science Education*, 12(3), 197-212.
- Zhao, J., & Jiang, Y. (2009). *A descriptive method for simulating a group knowledge building process*. Berlin: Springer.