CrossMark

Dynamics of undergraduate student generic problem-solving skills captured by a campus-wide study

Andis Klegeris[1] · Stephanie Barclay McKeown[2] ·
Heather Hurren[3] · Lindsay Joy Spielman[1] ·
Maegan Stuart[1] · Manpreet Bahniwal[1]

Abstract The ability to effectively problem solve is a highly valued competency expected of university graduates, independent of their area of study. Evaluation of problem-solving skill (PSS) development is hindered by a shortage of available tools for monitoring student progress and by lack of defined instructional strategies for development of these skills. Our research is aimed at addressing these problems. We have developed an evaluation tool, which we applied to study the dynamics of undergraduate student PSS. We tested first- and upper-year students

✉ Andis Klegeris
andis.klegeris@ubc.ca

Stephanie Barclay McKeown
stephanie.mckeown@ubc.ca

Heather Hurren
heather.hurren@ubc.ca

Lindsay Joy Spielman
lindsayspielman@live.ca

Maegan Stuart
maeganmjstuart123@hotmail.com

Manpreet Bahniwal
bahniwal.m@gmail.com

[1]   Department of Biology, University of British Columbia Okanagan Campus, 3187 University Way, Kelowna, BC V1V1V7, Canada

[2]   Planning and Institutional Research, University of British Columbia Okanagan Campus, 1138 Alumni Ave, Kelowna, BC V1V1V7, Canada

[3]   Centre for Teaching and Learning, University of British Columbia Okanagan Campus, 3333 University Way, Kelowna, BC V1V1V7, Canada

Springer

from 26 different courses (total enrollment of 2229 students). Overall improvement of PSS was detected for the first-year students over their first term of study. There were no significant differences between the PSS of first- and upper-year students, and no improvement was detected over a single term by measuring PSS in upper-year students. Only three courses were effective at facilitating PSS. Our data indicate that most of the standard lecture approaches do not develop undergraduate student PSS and that universities and individual instructors must take active steps to advance this critical skill set in university students.

Abbreviations
OECD   Organization for Economic Co-operation and Development
PBL   Problem-based learning
PSS   Problem-solving skills
PISA   Programme for International Student Assessment
PIAAC   Programme for the International Assessment of Adult Competencies

Introduction

There are a number of common skill sets that universities strive to develop in their undergraduate students. Recently, the term "Citizen Scholar" has been introduced by Arvanitakis and Hornsby (2016) to describe graduates that possess a set of key proficiencies to assist them in adapting to the ever-changing needs of society. These proficiencies include critical thinking, communication, problem-solving, research, and teamwork skills. Problem-solving skills (PSS) have recently become one of the most desirable competencies of graduating students due to the recognition that our world and its workforce are changing rapidly with global partnerships and technological advancements, which require maximum flexibility and adaptability (Bonikowska et al. 2008; Meo 2014). The Organization for Economic Co-operation and Development (OECD) believes that complex PSS are particularly in demand in fast-growing, highly skilled managerial, professional, and technical occupations (OECD 2014a). In addition to specific disciplinary skills and competencies, employers expect university graduates to have mastered a much broader range of employability skills, such as communication, leadership, problem-solving, and teamwork skills (Lowden et al. 2011). Our study concerns the development and measurement of generic or general PSS as opposed to discipline-specific PSS. There is ample research in the area of discipline-specific problem solving, especially in mathematics (Latterell 2003), chemistry (Gayon 2007), and physics (Adeyemo 2010), but scarce evidence is available with regard to the development and measurement of generic PSS (Greiff et al. 2014). University students themselves, as well as their future employers, university administrators, and funders highly value problem solving as one of the instrumental competencies expected of university graduates (Zúñiga-Vicente et al. 2011). Studies have been conducted by countries around the world to measure adult competencies, including problem solving, to assess current realities and prepare for better future economic well-being (Greiff et al. 2013; Hanushek et al. 2013). Results available from Statistics Canada's Adult Literacy and Life-Skills Survey and from the Programme for the International Assessment of Adult Competencies (PIAAC) by the OECD indicate that individual wealth, and subsequently, the social well-being of the country, is dependent on success in adult literacy and the ability to problem solve (Hanushek et al. 2013; OECD 2014b; Statistics Canada 2011).

Problem solving is defined as "an individual's capacity to use cognitive processes to confront and resolve real, cross-disciplinary situations in which the solution path is not immediately obvious and the literacy domains or curricular areas that might be applicable are not within a single domain of mathematics, science, or reading" (OECD 2004). The OECD's Programme for International Student Assessment (PISA) introduced problem-solving items for the first time in 2003 (Greiff et al. 2013). Their study examined the problem-solving abilities of 15-year-old students worldwide (OECD 2004). PISA findings indicated that good performance by a student in core school subjects does not necessarily mean that they are proficient in problem solving (OECD 2014b).

In a recent study, Statistics Canada assessed PSS in adult populations (16 to 65 years of age) from 11 different countries (Statistics Canada 2011). Their results pointed to peaks and plateaus of PSS development throughout the human life span; specifically, there were small incremental gains with each year of schooling, followed by a rapid rise in skill gains in the first few years of university education. However, after the initial years of university experience, there were no further improvements in PSS. Only those study participants who were employed in careers that required problem solving, such as managers and supervisors, saw continued growth in their problem-solving ability (Statistics Canada 2011). Demetriou (2003) also observed that an individual's processing functions level off at the end of the second or the beginning of the third decade of life.

Piaget identified complex problem solving as one of the key mental processes associated with child development (Inhelder and Piaget 1958). Based on these findings, William Perry attempted to shed light on the cognitive development of students, including their ability to use logical thinking and PSS, by proposing a nine-position model of intellectual development. These positions have been grouped into four progressive stages (Lang 2008; Perry 1970), although critics argue that an individual could be simultaneously experiencing several of these stages at any given time (Schommer-Aikins and Hutter 2002). PSS may appear during the second stage, termed multiplicity, where students begin to realize that there might be more than one solution to a problem. During the two subsequent stages (relativism and the commitment in relativism), students start using contextual knowledge, evaluate evidence, and integrate knowledge, which would enhance their PSS.

Most research on problem-solving ability of students performed thus far has been discipline specific. For example, Adeyemo (2010) looked at the discipline of physics and found a significant relationship between student academic ability and their PSS. A previous study by Chi et al. (1982) found that the subject-specific PSS level of students influenced how they approached a problem; expert students delved deep into the problem and considered many options, whereas novice students looked at surface-level solutions only. In a study involving mathematics students, Latterell (2003) investigated which teaching practices improved problem-solving ability of students and found that an enhanced mathematics curriculum provided for more problem-solving practice and led to better PSS over the traditionally taught mathematics curriculum. By reviewing the available problem-solving research data, Finney (2003) concluded that successful generic problem solving could not be taught but could be developed in disciplinary contexts, a conclusion also supported by Tricot and Sweller (2014). Woods (1998) described a program involving a series of small group, problem-based learning (PBL) workshop-style courses used to actively develop generic PSS. Cohen et al. (2000) proposed that high schools should be prioritizing the development of conceptual understanding of problem solving by students, as this process requires a significant amount of time that is not available in a university setting.

Evidence indicates that different methods of instruction could influence PSS development. Arvanitakis and Hornsby (2016) recommend that universities should reorient their instructional practices to better equip students as problem solvers. Cooper et al. (2008) examined the impact of collaborative group work on the development of PSS and found that students who worked in groups improved faster in their ability to problem solve when compared to students working individually. A set of generic problem-solving questions from the PISA database was used to demonstrate that undergraduate students exposed to the PBL method of instruction showed improvement in their generic PSS, which was not the case for students from several other courses that were exposed to standard lectures only (Klegeris et al. 2013a). The beneficial effect of PBL on student PSS is supported by a number of other studies (Lind 2016; Chetty 2014; Lian and He 2013; OECD 2014a; Woods 1998).

The current study examined the dynamics of generic PSS of first- and upper-year undergraduate students from several different programs on a relatively small university campus. In this study, 12 problem-solving items from PISA 2003 were adapted to create 2 different tests, which were administered to students at the beginning and the end of the academic term. This study design was applied to address the following research aims. First, we investigated whether there are significant differences between the problem-solving abilities of the first- and upper-year undergraduate students. Second, by comparing the PSS dynamics of students taking different individual courses across campus, we aimed to identify which instructional techniques, besides the PBL approach, may have beneficial effects on the PSS of undergraduate students. Third, by addressing the above two research objectives, we also wanted to examine, in a large sample of students, the utility of the generic problem-solving tests that were used as a tool for studies on the dynamics of PSS of university students.

Methods

Student recruitment

To address our research objectives, we planned to recruit approximately 1000 first- and upper-year students, which would be reflective of the 7700 undergraduate student population on our university campus. The two main criteria for selecting the courses were large class size and coverage of diverse areas of study on our campus. For some programs, third year is the final year of study, and in the majority of programs, most third- and fourth-year-level classes have a mix of third- and fourth-year students; therefore, a decision was made to compare the PSS of the first-year students with a combined population of third- and fourth-year students, which we have termed "the upper year of study."

For confidentiality reasons, we have coded individual courses and concealed the course titles. The courses have been coded (Y1-Hum-A, Y1-Hum-B, etc.) according to the year of study (Y1 vs. Y3/4) and the subject area. Humanity (Hum) courses include subjects such as English, art history, and theater; professional (Prof) courses may include engineering, nursing, and management; science (Sci) courses include subjects such as biology, chemistry, human kinetics, and physics; and social science (Soc) courses may include anthropology, economy, psychology, and geography. In some cases, courses had two sections (e.g., Y1-Prof-A1 and Y1-Prof-A2), where the same course content was presented to two separate groups of students by the same or different instructors.

This study was possible due to the voluntary participation of the students enrolled in the selected courses. Study protocols were approved by the Human Research Ethics Board of our university, and students willing to participate in the study voluntarily signed an appropriate consent form. The research team ensured that each student's consent and research data were hidden from the course instructors, who were offered access to data summaries only.

Study sample

Table 1 describes the courses and the student population participating in the study. Instructors from 26 different courses with a total enrollment of 2229 students allowed use of their class time for our study. Nine of the courses were at the first-year level, and 17 courses were at the third/fourth-year level. Class sizes varied considerably (12 to 265 students enrolled). Even though we attempted to select larger third- and fourth-year classes, there was a considerable difference in the class sizes between the first- and the third/fourth-year classes. Table 1 also shows the average course marks of all enrolled students obtained in the respective course, as well as their overall academic marks obtained in the term. The data support a common trend for Canadian universities in which average marks in the upper-year courses are typically higher than those in the first-year courses.

Since one of the objectives of this study was to compare the problem-solving ability of first- and upper-year students, for the data analysis we excluded participants (1) whose year level could not be identified due to missing information, (2) who were second-year students, and (3) who were third- and fourth-year students taking first-year classes. Students are not able to take upper-year classes during their first year of study. The number of eligible students shown in Table 1 represents the total number of students enrolled, after students were excluded based on the aforementioned criteria. Overall, almost half of the eligible students elected to participate in this study (953 students); 28.6% of the eligible students (580 students) completed the problem-solving test at both the beginning and the end of the term.

Problem-solving tests

The problem-solving tests and the method used to administer them have been described previously (Klegeris et al. 2013a). The 12 test items used in this study were selected from the item bank designed for the 2003 version of PISA developed by OECD. They were used to create two different tests, test 1 and test 2, each with the same number and type of items (OECD 2004). Each test was comprised of four problems (six questions) each. The maximum score for each of the two tests was 13 marks. Both test 1 and test 2 contained two "level 3" difficulty items (most challenging), three "level 2" difficulty items, and one "level 1" difficulty item (OECD 2004). Any further information related to the testing, including the marking key, is available to faculty interested in administering them by contacting the authors of this study.

A 15-min time limit for completion of the test by the university students was introduced (Klegeris et al. 2013a, b) in an attempt to standardize the testing conditions for participants completing the same tasks under the same testing conditions and time constraints (Brooks et al. 2003). An additional benefit of limiting the test time is to reduce the class time spent for this activity.

Table 1 Courses studied

Course code	Students enrolled	Average course mark (%)	Average term mark (%)	Eligible	Students writing beginning	Students writing end	Students writing both tests	Eligible students participating (%)	Writing both tests (%)
First-year courses									
Y1-Hum-A	85	76.8	74.8	64	26	21	21	40.6	24.7
Y1-Hum-B	110	70.9	68.3	98	26	19	18	27.6	16.4
Y1-Prof-A1	119	63.9	70.1	97	81	80	74	89.7	62.2
Y1-Prof-A2	90	63.2	70.9	89	55	45	44	62.9	48.9
Y1-Sci-A1	265	66.1	68.3	208	141	38	35	69.2	13.2
Y1-Sci-A2	247	75.1	77.8	243	29	18	16	12.8	6.5
Y1-Sci-B	160	81.6	76.1	130	125	72	70	97.7	43.8
Y1-Soc-A	116	67.5	69.2	93	42	20	20	45.2	17.2
Y1-Sci-J	189	76.3	78.2	182	10	7	7	5.5	3.7
Total year 1	1381	71.3	72.6	1204	535	320	305	45.7	22.1
Third- and fourth-year courses									
Y3/4-Hum-C	38	76.8	81.8	30	20	17	17	66.7	44.7
Y3/4-Prof-B	102	84.3	84.8	101	72	47	44	74.3	43.1
Y3/4-Sci-C	55	74.9	78.2	55	18	14	14	32.7	25.5
Y3/4-Sci-D	91	78.8	78.7	86	46	28	28	53.5	30.8
Y3/4-Sci-E	22	80.8	76.4	22	13	12	12	59.1	54.6
Y3/4-Sci-F	22	72.6	66.4	22	15	12	12	68.2	54.6
Y3/4-Sci-G	47	80.4	77.0	47	29	23	22	63.8	46.8
Y3/4-Sci-H	123	78.0	78.0	121	29	13	13	24.0	10.6
Y3/4-Sci-I	26	74.4	78.4	25	19	15	15	76.0	57.7
Y3/4-Soc-B	47	84.6	80.7	46	29	20	20	63.0	42.6
Y3/4-Soc-C	29	82.8	78.6	29	12	12	12	41.4	41.4
Y3/4-Soc-D	44	76.1	77.9	40	23	17	17	57.5	38.6

Table 1 (continued)

Course code	Students enrolled	Average course mark (%)	Average term mark (%)	Eligible	Students writing beginning	Students writing end	Students writing both tests	Eligible students participating (%)	Writing both tests (%)
Y3/4-Soc-E	102	88.2	82.7	101	27	21	18	29.7	17.7
Y3/4-Soc-F	40	75.3	75.5	39	18	10	10	46.2	25.0
Y3/4-Hum-D	22	83.8	83.9	22	9	8	7	45.5	31.8
Y3/4-Hum-E	12	89.2	80.6	12	5	5	5	41.7	41.7
Y3/4-Hum-F	26	82.3	82.1	26	11	9	9	42.3	34.6
Total year 3/4	848	80.2	78.9	824	395	283	275	48.9	32.4

To equate the two versions of the tests used in this study, we applied a linear equating method (see Online Resource 1) to determine that we were measuring the same content and cognitive processes across the two tests (Crocker and Algina 1986). We used a single group design with counterbalancing (Ryan and Brockmann 2011). Problem-solving tests 1 and 2 were administered to students in a random manner at the beginning of the term (September), so that 494 students wrote test 1 first and 458 students wrote test 2 first. At the end of the term (end of November/beginning of December), individual students completed the other of the two tests, test 2 if they wrote test 1 in September and vice versa. As mentioned previously, the tests were designed to be parallel in content by selecting the same number of items and the same level of difficulty as identified by PISA researchers. Since the same students wrote both version of the test, any differences in scores could be attributed to differential difficulty of the tests (Ryan and Brockmann 2011). This way, students participating at the beginning and at the end of the term wrote two tests that were different, yet comparable in difficulty.

All tests were marked in a blinded manner by randomly shuffling coded tests into a single pile. Marking was done by a group of volunteers according to a pre-determined answer key. Each volunteer was responsible for marking the same one question on all tests, thus avoiding inter-rater variability. The student test scores were linked to their final academic grades earned during the term when our study was conducted (both in the course studied and average term grade for all courses taken by the student) by staff members in the university registrar's office.

Results

The unadjusted results from tests 1 and 2 are reported in Table 2 by year level. Prior to equating, the raw mean results for the two tests were compared, by year level, using a Student's t test that indicated a statistically significant difference between the scores achieved in test 1 and test 2 ($p < 0.01$) for both the first- and upper-year students. The results indicated that test 2 was statistically significantly more difficult than test 1.

Next, using Eq. 1 (see Online Resource 1), the original scores on test 1 were adjusted using a linear equating procedure to transform the original score to a score that would be equal to the score on test 2. These adjusted test 1 scores were used for all subsequent analyses. Figure 1a shows that there was a statistically significant increase in problem-solving abilities of the first-year students after their first term of study at the university (13.7% increase, $p < 0.01$), while the 2.9% increase in the test score for the upper-year students was not statistically significant ($p = 0.4$). Comparison of the first- and upper-year student scores at the beginning of the term showed that the upper-year student scores were better than the first-year student scores by

Table 2 Original test scores (out of 13) by year level

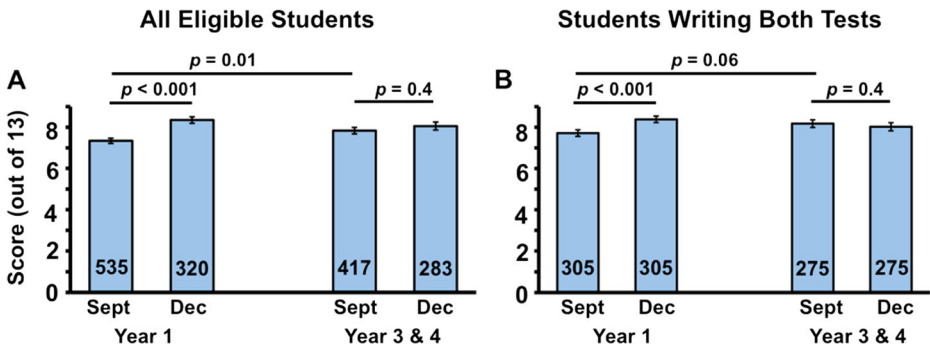|	Mean	Median	Standard deviation
First-year students			
Test 1	8.6	9.0	2.8
Test 2	7.7	8.0	2.9
Third- and fourth-year students			
Test 1	8.6	9.0	3.0
Test 2	7.9	8.0	2.9

Fig. 1 Average problem-solving test scores obtained by the first- and upper-year students at the beginning and end of the term

6.7%, which was also statistically significant ($p = 0.01$); however, this comparison should be viewed as preliminary data only since these data were collected from two different groups of students. Our follow-up longitudinal studies will be aimed at further examining these observations.

Figure 1a also illustrates that there were more students who completed the test at the start of the term than at the end (535 compared with 320 for the first-year courses and 417 compared with 283 for the upper-year courses). The average course marks of the students writing only the beginning of the term test were significantly lower compared to students writing tests both at the beginning and end of the term. For example, for the first-year students who wrote the first test only, average course mark was $67.8 \pm 1.0\%$, while students writing both tests achieved an average mark of $72.7 \pm 0.8\%$ ($p < 0.001$; Student's t test). These observations indicate that these two groups of students may have intrinsic differences in their abilities and represent two different student populations; therefore, we re-calculated the average problem-solving test scores of all first- and upper-year students by using data collected from eligible students who wrote both tests (305 first-year students and 275 upper-year students).

Figure 1b illustrates that removing data from students who wrote just one test led to a decrease in the measured improvements of problem-solving test scores and statistical significance levels. The results of the Student's t test indicated that the increase in problem-solving abilities of the first-year students after their first term of study remained statistically significant, but the improvement in test scores was reduced (8.7% increase, $p < 0.01$). An effect size of 0.24 was calculated using Eq. 2 (see Online Resource 1), where the difference between the test means was divided by the pooled standard deviation. Cohen (1988) suggested to interpret the values of d as small ($d = 0.2$), medium ($d = 0.5$), or large ($d = 0.8$). Following this interpretation guideline, a small effect size of 0.24 indicates that students performed better on their end-of-term tests compared to the beginning of the term tests by about a quarter of a standard deviation (Cohen 1988); however, more recently, Lakens (2013) has suggested that an effect size should be interpreted in relation to other similar reported effects. There was a 1.8% decrease in the test score for the upper-year students, but this effect was not statistically significant ($p = 0.4$). Comparison of the first- and upper-year student scores at the beginning of the term showed a 6.0% improvement ($p = 0.06$).

In the next step, we analyzed data obtained from individual courses. We used scores from the students completing both tests for all the subsequent analyses described in the

following section. Only courses in which 10 or more students wrote both tests were included in the analysis. Figure 2 shows the data obtained from the 22 courses that met this criterion, out of the 26 courses that participated in the study. Figure 2a presents the
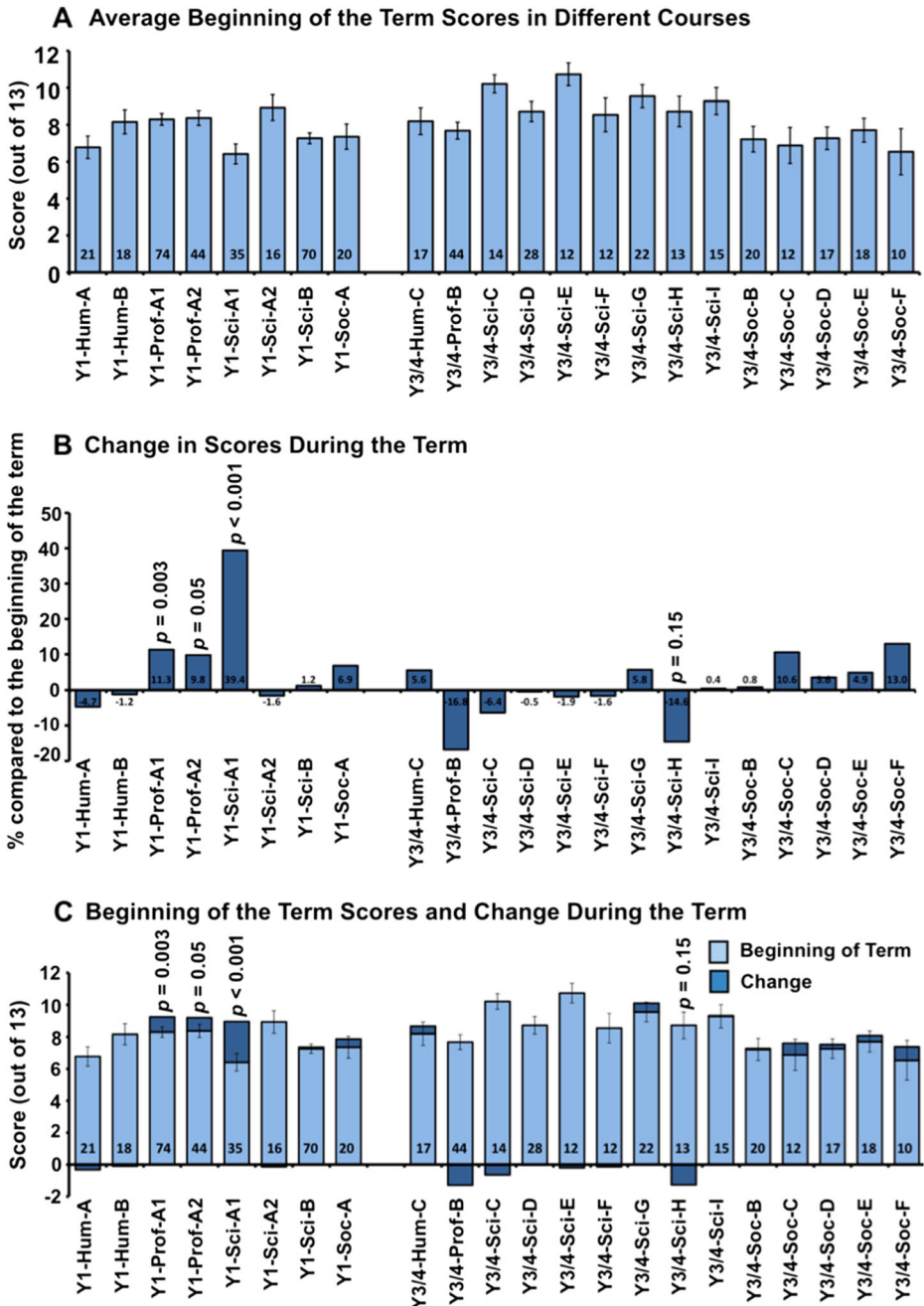

Fig. 2 Comparison of the problem-solving scores obtained by students from different first- and upper-year courses

average problem-solving test scores obtained by students taking each of these courses at the beginning of the term; Fig. 2b shows the change in average test scores during the term, which is expressed as percent change; and Fig. 2c combines the previous two figures and shows absolute values of both the beginning of the term score and the change in the test scores.

Figure 2a illustrates that there were large differences in average scores obtained by students from different courses during the beginning of the term testing. Thus, the highest and lowest averages for the first-year course were 8.9 and 6.4 out of 13, respectively. Corresponding values for the upper-year courses scores were 10.7 and 6.5 out of 13. A one-way ANOVA was performed on the average beginning-of-term scores for first- and final-year courses, using a Bonferroni correction to determine which specific courses differ from the rest. Within the sample of the eight first-year courses, there was a statistically significant difference among the course means at the $p < 0.01$ level [$F(7, 297) = 23.54$, $p = 0.003$]; however, the results of the post hoc analysis identified that there was only one course (Y1-Sci-A1) where the mean differed statistically significantly from several other courses. As for the final-year courses, again the results of the ANOVA indicated statistically significant differences among the course means at the $p < 0.01$ level [$F(13, 253) = 22.92$, $p = 0.001$], yet the results of the post hoc Bonferroni test indicated that differences between the individual course means were not statistically significant.

Figure 2b shows that PSS of students in five out of the eight first-year courses improved. There were small decreases in three courses, which were not statistically significant according to the paired Student's t test used for all subsequent statistical analyses. Of the five courses with improved scores, in the following two cases, this change was highly significant: $p = 0.003$ for Y1-Prof-A1 and $p < 0.001$ for Y1-Sci-A1. The improvement for the third course (Y1-Prof-A2) showed a very strong trend towards significance $p = 0.05$. Figure 2b also illustrates that PSS improvements recorded in three different courses were above the average value of 7.8% with one of the courses (Y1-Sci-A1) showing a very large 39.4% improvement. In order to exclude the possibility that the significant gains in PSS of all the first-year students depended exclusively on this one large course with 35 students writing both tests, we performed comparison of the beginning and end of the term test scores of first-year students after removing all data obtained from this course. The increase in PSS of first-year students remained statistically significant (270 students, 5.6% improvement, $p = 0.01$; data not shown).

For the upper-year courses, no statistically significant changes were observed with $p = 0.15$ being the lowest p value achieved and most other p values exceeding 0.2. Nine out of 14 upper-year courses showed small increases in test scores. There were small decreases in test scores for the other five courses, which were not statistically significant.

We performed a similar analysis of students' test scores after grouping all eligible students who completed both tests according to their majors (programs of study). Since most students in our institution enroll in various majors only after the second year of their study, such analysis was possible for the upper-year students only. Figure 3 shows the data from those majors where 10 or more students completed both tests. Of the 11 majors, in 6, there was a small increase, and in 5, there was a small decrease in the test scores; none of which were statistically significant according to the paired Student's t test. The p values that were less than 0.2 are shown in Fig. 2b, c and 3b.
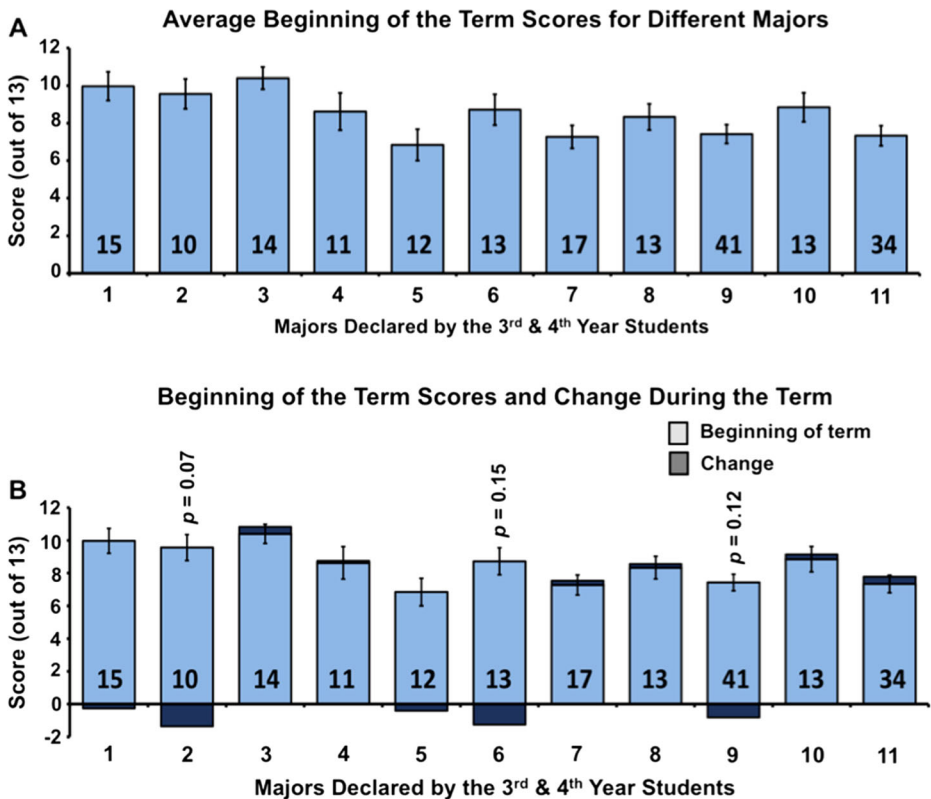
Fig. 3 Comparison of the problem-solving scores obtained by upper-year students from different majors

Discussion

To the best of our knowledge, this is the first campus-wide study of the dynamics of generic PSS of undergraduate students. We achieved our initial target of enrollment of 1000 students in the study by recruiting 1176 students who signed consent forms. Data from a total of 953 students were eligible for the study. One of the first observations that we made was a substantially lower student participation rate at the end of the term compared to the beginning of the term. We hypothesized that the more motivated, and possibly higher achieving, students formed a distinct subpopulation of students attending courses at the end of the term. This could lead to an overestimated gain of PSS if this group also demonstrated higher ability to master these skills.

Our hypothesis of distinct subgroups of students was supported by the analysis of academic course grades of these students; they were significantly higher for the group of students completing both tests at the beginning and end of the term compared to the group of students completing the test at the beginning of the term only. Therefore, for the purposes of this discussion, we will only consider data obtained from the 580 students who completed both tests. It is important to note that all problem-solving tests were marked in a blinded manner to avoid marking bias (Klegeris and Hurren 2011). There are several notable conclusions and trends that could be drawn from the data obtained, which relate to the three initial study objectives.

Problem-solving abilities of the first- and upper-year undergraduate students

A statistically significant 8.7% improvement in problem-solving ability was detected for the first-year students during their first term of study at the university. No improvement could be observed for the upper-year students with the 1.8% downward trend not being statistically significant. Comparison of the test scores of first- and upper-year students showed a 6.0% improvement with a strong trend towards significance ($p = 0.06$), which may indicate that university experience over 3 to 4 years of study has beneficial effects on PSS of students; however, this conclusion will require future confirmatory studies since our current data were obtained from two different, non-overlapping groups of students who may have differing intrinsic problem-solving abilities. Overall, the lack of any significant downward trajectory in the PSS of undergraduate students indicates that their university experience on our campus does not negatively impact their PSS.

Our main conclusions with regard to the dynamics of student problem-solving abilities are in line with the report published by Statistics Canada, which indicates that the initial years of study at the university are beneficial with respect to PSS, while during most of adulthood, these skills seem to fluctuate, unless specifically developed through ongoing job requirements (Statistics Canada 2011). The 8.7% improvement observed in this large ($N = 305$) group of the first-year students was less than the 13% increase previously recorded in a single third-year course ($N = 47$) of students exposed to PBL (Klegeris et al. 2013a).

The 5–39% changes in PSS recorded by testing various individual groups of students observed in this and other studies (Klegeris et al. 2013a, b) are consistent with data obtained in several different countries by the OECD study. For example, the OECD study recorded 7–34% improvement in PSS over 25 years of formal education in the age group of 16–65-year olds, which depended on the country of study, and 9–12% decline in PSS of adults as they age from 26 to 65 years, which depended on the level of education and the type of occupation (Statistics Canada 2011).

Methods of instruction having beneficial effects on problem-solving skills of undergraduate students

One of the goals of this study was identification of courses, and hence instructional techniques, which could facilitate development of PSS of students. Our search covered prototypical courses from the diverse programs and disciplines offered on our campus (see Table 1); therefore, none of the courses studied had any performance-based enrollment requirements such as honors standing of students or certain minimal level of test scores. Analysis of our data showed that, first, there were significant differences in the test scores obtained by students from different courses during the beginning of the term testing (see Fig. 2a), indicating that different student subpopulations could have different problem-solving abilities. There are no obvious differences in the student populations enrolling into the different programs on our campus, except that certain applied science courses tend to have a very high percentage of male students, while nursing courses, for example, have higher than average female representation. We did not study the effects of gender on PSS of students, but previous research has failed to identify gender differences in PSS of adults (Statistics Canada 2011). The causes for such differences between the PSS of groups of students attending various classes will require further study.

According to our data, there were no courses that caused a significant decrease in student problem-solving ability. Our study failed to identify upper-year courses that would lead to significant improvement in student PSS. Grouping the upper-year students according to the majors they had declared also did not lead to identification of a program that benefited development of PSS. Comparison between the first-year and upper-year students in various majors was not performed since most students at our university are only able to declare their major after completing their first year of study. However, there were three first-year courses where the improvement during the first term of study was either statistically significant or had a very strong trend towards significance (see Fig. 2b). The causes for the variations in the dynamics of PSS test scores among the different courses studied remain unknown, but differences in the course content between the classes investigated could be one of the reasons for such observed diversity.

Another aspect of the courses that could have been responsible for the observed variation in the dynamics of PSS is the different instructional techniques used in various courses. For example, use of PBL, case studies, or team-based learning could affect the PSS of students (Klegeris et al. 2013a). We did not use any objective or quantitative measures to characterize the diversity of instructional styles, but we did briefly interview all participating instructors on the instructional techniques they used in the courses we studied. In most courses, other than the ones described in the following paragraph, the methodologies used could be described as a traditional lecture-style course content delivery, enhanced by in-class discussion.

The following two of the first-year courses studied were different sections of a course for a professional program that was instructed by the same professor: Y1-Prof-A1, 11.3% improvement, $N = 74$, $p = 0.003$, and Y1-Prof-A2, 9.8% improvement, $N = 44$, $p = 0.05$. The fact that both sections of this course saw significant PSS improvement indicates strongly that the instruction both these groups of students received during their first term of study helped with their generic PSS. An interview with the course instructor revealed that the Y1-Prof-A course had a classical lecture component (2 h per week) and a tutorial component (2 h per week), where students were split into groups of 20 and worked individually on open-ended problems. In addition, students were required to work individually on related sets of problems by use of an online tool, which also evaluated their performance in this activity. Therefore, the instructional techniques used in this course may have contributed to the improvement of PSS of the students taking it, but this hypothesis will require further investigation.

The largest improvement of PSS observed in this or any of previous studies was in Y1-Sci-A1 (39.4% improvement, $N = 35$, $p < 0.001$). Notably, the other section of this course instructed by a different professor showed no improvement in PSS, Y1-Sci-A2, 1.6% decrease, $N = 16$, $p > 0.2$. It is important to note that the increase in the test scores observed for Y1-Sci-A1 brought the average score for this course at the end of the term to a level that was comparable with several other science courses including Y1-Sci-A2, while at the beginning of the term, this group of students obtained the lowest score of all courses studied. The most likely reasons for such dramatic disparity between the two sections of the same course include varying instructional styles between the two professors instructing the two sections of the course, or possibly some external factor interfering with the students from Y1-Sci-A1 during the beginning of the term test, which could have led to artificially low average test score. However, interviews with the instructors revealed no notable differences between the instructional styles of the professors; both sections were instructed mainly through standard lecturing, both professors were very experienced and had instructed this course for several years, and there was no obvious difference in student enrollment into the two sections of this course.

Therefore, the significant improvement of PSS of Y1-Sci-A1 students observed in this study will require future confirmation.

Utility of the generic problem-solving test

Data obtained by this study of a large student population corresponded to the previous reports by Klegeris et al. (2013a, b), showing that the generic problem-solving tests used could detect changes in PSS of students after only one term of study (3 months). After adjusting for differences in the difficulty level of the test 1 and test 2 scores, a statistically significant improvement in PSS was detected for a large population of the first-year students ($N = 305$) as well as for individual groups of students enrolled in two different courses ($N = 35$ and $N = 74$). Thus far, statistical significance for individual courses has been achieved in those cases where the improvement exceeded 11%. Strong trends towards significance have been observed with effects ranging from 9 to 10% (see Fig. 2b; Klegeris et al. 2013b). It was not possible to detect differences in PSS of students exposed to PBL for less than 3 months (Klegeris et al. 2013b); therefore, data collected thus far indicate that test subjects in groups of up to 100 need to be exposed to the beneficial environment for several months and achieve >10% improvement in the test scores for the changes to be detected by the tests we used in this study. In a larger group of students (>200), even a 5.6% average increase in PSS was detected by the test system used. Our study supports the utility of this relatively simple technique in studies measuring the generic PSS of students, as well as adult populations over 18 years of age. Noteworthy advantages of this test system include the following: (1) it is freely available, (2) it takes only two 15-min periods to administer, and (3) use of two comparable tests in a crossover fashion allows short-term studies of the dynamics of PSS.

Significance and advancement of problem-solving skills

We concur with a number of authors who consider development of PSS as one of the most important, if not the ultimate, goal of any educational system (Adeyemo 2010; Gagne 1985; Meo 2014). This realization has come to the forefront of the educational literature during recent years due to the fact that the emphasis in education has been gradually shifting from equipping students with highly codified discipline- and subject-specific skills to empowering them to confront and overcome complex, non-routine cognitive challenges common to various work environments. Recent decades have seen a rapid increase in the jobs that require high levels of creative PSS (OECD 2014a). This study was designed to address the two main challenges that hinder development of PSS of adults, including university students, which are (1) the lack of widely available and accepted tools for assessing PSS and (2) the shortage of knowledge about instructional techniques and curricular activities that would have reliable and reproducible beneficial effects on the PSS of adult learners (Greiff et al. 2014). This study demonstrated the utility of a simple testing approach that could help tackle the first of the challenges. With regard to the second challenge, there were no immediate answers given by this study as to the best approaches that could be used to develop PSS. Instead, we identified a professional program on our campus, which is very successful with developing the generic PSS in a large cohort of their first-year students. We hope that our future research, focused on instructional techniques within this program, will help pinpoint the factors that have beneficial

effects on the problem-solving abilities of these students. We currently do not have data on the PSS of the upper-year students from this program, which is something that we are planning to address in our follow-up studies.

Universities have traditionally been founded upon the timeworn principle that the sole purpose of an educational institution is to impart discipline-specific knowledge upon its students. In today's society, however, where knowledge can be gained through numerous readily accessible sources such as, journals, newspapers, and a variety of internet-based sources, universities no longer control the dissemination of knowledge as they once did (Arvanitakis and Hornsby 2016; Sireteanu and Bedrule-Grigoruta 2007). Furthermore, the dynamic work places of this era demand well-rounded employees who not only demonstrate the industry-specific skill set and competencies but who are also able to work across multiple geographical and cultural locations, interpret vast amounts of data, and integrate information across several disciplines (Lind 2016). Pool and Sewell (2007) asserted that an individual's ability to both obtain and retain employment in this dynamic occupational culture is reliant on a number of different skills and attributes that extend beyond subject-specific knowledge and abilities. These essential attributes include imagination/creativity, adaptability, willingness to learn, ability to work both independently and as part of a team, effective communication, organizational abilities, emotional intelligence, and self-confidence (Pool and Sewell 2007).

Since one of the paramount roles of universities is to prepare students for the post-graduate work force, universities should make the effort to adapt to the changing workplace climate, by fostering the development of functional, desirable, and discipline-transferable skills in their students. Arvanitakis and Hornsby (2016) pleaded that universities must "redefine their role" or risk becoming redundant, and instead of focusing on solely providing students with discipline-specific knowledge, universities should aim to improve the development of student skills such as creativity, resilience, design thinking, and collaboration, all of which are components of effective PSS. Although modifying the university structure and culture to achieve this objective would require a great deal of work, this is not by any means an impossible task, as this idea has already been adopted and implemented by some educational institutions across the world. The University of Luton in England has taken tremendous steps in revising its university structure and curriculum to ensure that its students not only learn discipline-specific content, but instead, the curriculum emphasizes the development of trans-ferable skills that are necessary for future employment and lifelong learning (Fallows and Steven 2000). As a result, University of Luton has a greater than average proportion of its graduates entering into permanent employment shortly after graduation (Fallows and Steven 2000).

Previous research has already identified several different instructional strategies, which can have a beneficial effect on the development of PSS in students. PISA 2012 results show that the structure of educational curricula and quality of classroom teachers make a difference in imparting PSS (OECD 2014b). It has been demonstrated that creating a safe classroom learning environment, by integrating more opportunities for one-on-one interactions between students and their professor, asking for class feedback, and encouraging student questions, is essential to create a classroom setting where critical thinking skills can truly be developed (Brenner 2016). It has also been argued that deeper and broader learning can be achieved when instructors pose open-ended questions that ask students to link discipline-specific knowledge with everyday experiences, thus promoting the development of broader critical thinking skills (Brenner 2016; Zivkovic 2016). Engaging students in active learning, through

practices of discussion, debate, and PBL, promotes processes of higher-order thinking, such as analysis, as well as knowledge integration and evaluation, and thereby helps to foster the development of generic PSS (Bonwell and Eison 1991; Zivkovic 2016).

Cooper et al. (2008) demonstrated that collaborative group work could improve PSS of students. PBL has been identified as a technique that has reliable positive effects on student PSS (Chetty 2014; Duch et al. 2001; Gayon 2007; Klegeris et al. 2013a, b; Lian and He 2013; Woods 1998). Anderson et al. (2008) found that using PBL cases in an online environment translated to higher scores on authentic assessments using problem solving. Thus, PBL involving group-based activities is emerging as one of the possible instructional techniques, which can be used to advance PSS of undergraduate students and other adult learners. Recent research has also demonstrated that the traditional small-group PBL technique can be modified for use with large groups of students without engaging additional tutors. This approach is more cost-effective, yet remains a sound instructional method to enhance PSS and improve student satisfaction (Klegeris and Hurren 2011; Tune et al. 2013; Walsh 2013; Woods 1998).

Our data confirmed overall improvement of PSS of first-year students over their first term of study at our university. However, students from only three out of the eight first-year courses that participated in this study showed either a trend towards significance or significant improvement in PSS. Furthermore, none of the 14 upper-year courses participating in this research had any effect on PSS of students. These data indicate that while some university instructional techniques on our campus may be beneficial in developing PSS skills of undergraduate students, other classroom strategies could be improved in order to foster the development of undergraduate PSS skills. Without further studies, our data cannot be extrapolated to student populations outside our university campus, and we are actively looking for collaborators from other universities who could carry out such studies. Nonetheless, our results clearly demonstrate that only select courses (approximately 10% of the courses studied thus far on our campus) have significant beneficial effects on student PSS. If these observations are supported by studies in other universities, it could be concluded that most of the standard lecture approaches are not effective at improving PSS of students and that in order to advance this important skill set, active steps need to be taken by universities as well as individual instructors.

Several university programs have been developed and implemented to improve the teaching and learning of important transferable skills that are essential for survival in the post-graduate world (Fallows and Steven 2000; Ameredes et al. 2015). For example, the University of Texas Medical Branch has developed and implemented a multidisciplinary translational team model for training researchers at the graduate student level. The multidisciplinary translational team consists of researchers and medical professionals at all career stages, who use face-to-face meetings to allow for regular interactions with students seeking mentorship, to promote the discussion of ideas and challenges, and to help students find creative solutions to research-related problems. This program has produced beneficial outcomes for their graduate students, such as the improvement of translational research competencies and teamwork skills, as well as an increased number of collaborative research publications, which assists students in gaining and retaining employment in the future (Ameredes et al. 2015). Other universities have adopted and applied teaching exercises that focus on developing skills that are necessary for life-long learning, including experiential exercises, PBL, case studies, and continual invitation for student engagement through classroom discussion (Fallows and Steven 2000).

Conclusion

There are many instructors, who constantly seek out new and improved methods for teaching subject-specific knowledge, with the purpose of simultaneously improving other real-life applicable skills of their students. However, university-wide programs aimed at developing such skills are not common practice at the undergraduate level, and a vast majority of universities and university programs still operate as archaic pedagogical systems of knowledge dissemination. Therefore, the creation of university programs that foster the development of transferable skills, such as critical thinking, teamwork skills, and PSS, as well as tangible incentives motivating professors and instructors to implement instructional techniques, which have been proven to achieve the aforementioned goals, are necessary if universities wish to remain relevant in the shifting educational and workplace culture (Arvanitakis and Hornsby 2016). Moreover, by advancing essential modern-life skills of students, such as PSS, and documenting their progress in this important cognitive domain, universities will be able to demonstrate the benefits of higher education, which go beyond teaching subject-specific information, and enable them to stay competitive in the era of increasingly diverse educational opportunities for students.

References

Adeyemo, S. A. (2010). Students' ability level and their competence in problem-solving task in physics. *Int J Edu Res Technol, 1*(2), 35–47.

Ameredes, B. T., Hellmich, M. R., Cestone, C. M., Wooten, K. C., Ottenbacher, K. J., Chonmaitree, T., Anderson, K. E., & Braisier, A. R. (2015). The multidisciplinary translational team (MTT) model for training and development of translational research investigators. *Clin Trans Sci, 8*(5), 533–541.

Anderson, W. L., Mitchell, S. M., & Osgood, M. P. (2008). Gauging the gaps in student problem-solving skills: assessment of individual and group use of problem-solving strategies using online discussions. *CBE Life Sci Edu, 7*(2), 254–262.

Arvanitakis, J., & Hornsby, D. J. (2016). Are universities redundant? In J. Arvanitakis & D. J. Hornsby (Eds.), *Universities, the citizen scholar and the future of higher education* (pp. 7–20). New York, NY: Palgrave Macmillan.

Bonikowska, A., Green, D. A., & Riddell, W. C. (2008). *Literacy and the labour market: cognitive skills and immigrant earnings (catalogue no. 89–552-M no. 020). Survey.* Ottawa, CA: International Adult Literacy.

Bonwell, C. C., & Eison, J. A. (1991). *Active learning: creating excitement in the classroom. ASHE-ERIC higher education report no. 1.* Washington, DC: George Washington University.

Brenner, E. (2016). Changing mindsets: moving from the acceptance of facts to critical thinking. In J. Arvanitakis & D. J. Hornsby (Eds.), *Universities, the citizen scholar and the future of higher education* (pp. 155–171). New York, NY: Palgrave Macmillan.

Brooks, T. E., Case, B. J., & Young, M. J. (2003). Assessment report: timed versus untimed testing conditions and student performance. Retrieved from http://images.pearsonassessments.com/images/tmrs/tmrs_rg/TimedUntimed.pdf?WT.mc_id=TMRS_Timed_Versus_Untimed_Testing. Accessed 28 Oct 2016.

Chetty, N. (2014). The first year augmented programme in physics: a trend towards improved student performance. *S Afr J Sci, 110*(1/2), 1–9.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7–76). Hillsdale, NJ: Erlbaum.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J., Kennedy-Justice, M., Pai, S., Torres, C., Toomey, R., DePierro, E., & Garafalo, F. (2000). Encouraging meaningful quantitative problem solving. *J Chem Educ, 77*(9), 1166–1173.

Cooper, M. M., Cox Jr., C. T., Nammouz, M., & Case, E. (2008). An assessment of the effect of collaborative groups on students' problem-solving strategies and abilities. *J Chem Educ, 85*(6), 866–873.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* New York, NY: Holt, Rinehart, and Winston.

Demetriou, A. (2003). Mind, self, and personality: dynamic interactions from late childhood to early adulthood. *J Adult Dev, 10*(3), 151–171.

Duch, B. J., Groh, S. E., & Allen, D. E. (2001). *The power of problem-based learning: a practical "how to" for teaching undergraduate courses in any discipline.* Sterling, VA: Stylus Publishing.

Fallows, S., & Steven, C. (2000). Building employability skills into the higher education curriculum: a university-wide initiative. *Educ Train, 42*(2), 75–82.

Finney, R. (2003). Research in problem-solving: improving the progression from novice to expert. Retrieved from http://www.colorado.edu/physics/phys4810/phys4810_fa06/4810_readings/finney.pdf. Accessed 28 Oct 2016.

Gagne, R. M. (1985). *The conditions of learning* (4th ed.). New York, NY: Holt, Rinehart &Winston.

Gayon, E. E. P. (2007). The problem-solving ability of high school chemistry students and its implications in redefining chemistry education. Paper presented at the International Conference on Science and Mathematics Education, Ponang, Malaysia.

Greiff, S., Holt, D. V., & Funke, J. (2013). Perspectives on problem solving in educational assessment: analytical, interactive, and collaborative problem solving. *J Probl Solv, 5*(2), 71–91.

Greiff, S., Wüstenberg, S., Csapó, B., Demetriou, A., Hautamäki, J., Graesser, A., & Martin, R. (2014). Domain-general problem solving skills and education in the twenty-first century. *Edu Res Rev, 13*, 74–83.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2013). Returns to skills around the world: evidence from PIAAC. National Bureau of Economic Research. http://www.nber.org/papers/w19762.pdf. Accessed 28 Oct 2016.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York, NY: Basic Books.

Klegeris, A., & Hurren, H. (2011). Impact of problem-based learning in a large classroom setting: student perception and problem-solving skills. *Adv Physiol Educ, 35*(4), 408–415.

Klegeris, A., Bahniwal, M., & Hurren, H. (2013a). Improvements in generic problem-solving abilities of students by use of tutor-less problem-based learning in a large classroom setting. *CBE Life Sci Education, 12*(1), 73–79.

Klegeris, A., Bahniwal, M., & Hurren, H. (2013b). Tutor-less problem-based learning in a large classroom setting significantly enhances generic problem-solving skills of undergraduate students. Paper presented at the 6th International Conference of Education, Research and Innovation (ICERI). Seville, Spain: IATED

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology: Cognition, 4*, 1–12.

Lang, J. M. (2008). *On course: a week-by-week guide to your first semester of college teaching.* Cambridge, MA: Harvard University Press.

Latterell, C. (2003). Testing the problem-solving skills of students in an NCTM-oriented curriculum. *Math Educ, 13*(1), 5–14.

Lian, J., & He, F. (2013). Improved performance of students instructed in a hybrid PBL format. *Biochem Mol Biol Educ, 41*(1), 5–10.

Lind, J. M. (2016). Medical education: training for the desirable traits in past, present and future doctors? In J. Arvanitakis & D. J. Hornsby (Eds.), *Universities, the citizen scholar and the future of higher education* (pp. 172–186). New York, NY: Palgrave Macmillan.

Lowden, K., Hall, S., Elliot, D. & Lewin, J. (2011). Employer's perceptions of the employability skills of the new graduates. Edge/The SCRE Centre, University of Glasgow. http://www.educationandemployers.org/wp-content/uploads/2014/06/employability_skills_as_pdf_-_final_online_version.pdf. Accessed 28 Oct 2016.

Meo, S. A. (2014). Undergraduate medical student's perceptions on traditional and problem based curricula: pilot study. *J Pak Med Assoc, 64*(7), 775–779.

Organisation for Economic Co-operation and Development, OECD (2004). *Problem solving for tomorrow's world: first measures of cross-curricular competencies from PISA 2003.* Paris: OECD https://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/34009000.pdf. Accessed October 28 2016.

Organisation for Economic Co-operation and Development, OECD (2012). *Assessment of Higher Education Learning Outcomes (AHELO) feasibility study report (volume I)*. Paris: OECD http://www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume1.pdf. Accessed 28 October 2016.

Organisation for Economic Co-operation and Development, OECD (2014a). *PISA 2012 results: creative problem solving: student' skills in tackling real-life problems (volume V)*. Paris: OECD http://www.oecd.org/pisa/keyfindings/pisa-2012-results-volume-v.htm. Accessed 28 October 2016.

Organisation for Economic Co-operation and Development, OECD (2014b). *PISA 2012 results in focus: what 15-year olds know and what they can do with what they know*. Paris: OECD https://www.oecd.org/pisa/keyfindings/pisa-2012-results-volume-I.pdf. Accessed 28 October 2016.

Perry Jr., W. G. (1970). *Forms of intellectual and ethical development in the college years: a scheme*. New York, NY: Holt, Rinehart, and Winston.

Pool, L. D., & Sewell, P. (2007). The key to employability: developing a practical model of graduate employability. *Educ Train, 49*(4), 277–289.

Ryan, J., & Brockmann, F. (2011). A practitioner's introduction to equating: with primers on classical test theory and item response theory. Council of Chief State School Officers. http://www.ccsso.org/Documents/2008/A_Practritioners_Introduction_To_2008.pdf. Accessed 28 Oct 2016.

Schommer-Aikins, M., & Hutter, R. (2002). Epistemological beliefs and thinking about everyday controversial issues. *J Psychol, 136*(1), 5–20.

Sireteanu, N. A., & Bedrule-Grigoruta, M. V. (2007). Perspectives of knowledge management in universities. *Soc Sci Res Network (SSRN) Electron J*. doi:10.2139/ssrn.1029990 .Accessed 28 October 2016

Statistics Canada. (2011). Learning a living: first results of the adult literacy and life skills survey, Chapter 5, Adult problem solving skills. http://www.oecd.org/education/innovation-education/34867438.pdf. Accessed 28 Oct 2016.

Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching complex skills does not work. *Educ Psychol Rev, 26*(2), 265–283.

Tune, J. D., Sturek, M., & Basile, D. P. (2013). Flipped classroom model improves graduate student performance in cardiovascular, respiratory, and renal physiology. *Adv Physiol Educ, 37*(4), 316–320.

Walsh, K. (2013). Tutor-less problem-based learning in a large classroom setting—could it also save costs? *CBE Life Sci Edu, 12*(3), 315.

Woods, D. R. (1998). The McMaster problem solving (MPS) program. Retrieved from http://chemeng.mcmaster.ca/mcmaster-problem-solving-mps-program. Accessed 28 Oct 2016.

Zivkovic, S. (2016). A model of critical thinking as an important attribute for success in the twenty-first century. *GlobELT, 2016*(232), 102–108.

Zúñiga-Vicente, J. A., González, A. B., & Román, C. P. (2011). How similar or different are the generic skills among different areas of knowledge? A comparative analysis in the Spanish higher education. Paper presented at the International Conference on Education and New Learning Technologies, Barcelona, Spain.