PERCEPTIONS OF DIFFICULTY

Leigh N WOOD
University of Technology, Sydney
Broadway, Australia, 2007
leigh.wood@uts.edu.au

Geoffrey H SMITH
University of Technology, Sydney
Broadway, Australia, 2007
gsmith@maths.uts.edu.au

ABSTRACT

Seventy students from a first semester calculus course ranked 8 mathematics tasks as to perceived difficulty before attempting these tasks and actual difficulty after completing the tasks. Students also completed two examinations, one based on facts and procedures and the other based on applications and concepts. The tasks were designed to fit into a taxonomy of mathematical skills.

We have found that students perceive questions to be difficult for a number of reasons. In general, questions requiring conceptual understanding are regarded as more difficult that those which require factual recall or the use of routine procedures. There was not a strong link between familiarity with the question type and ranking of difficulty. Students were sufficiently familiar with the some types of question to be able to perceive inherent difficulties, such as a complex differentiation.

We found that in five out of eight cases, students' perceptions of the difficulty did not change after they had done the task. In one case they found the question to be more difficult than expected and in two cases to be easier. It is not clear to us why students found one question to be more difficult than expected. It may be that some of the complexities (such as the use of the intermediate value theorem) were not immediately apparent. It is also significant that NESB students rated this question as easier than ESB students. This was the case both before and after attempting the question. Student comments are also presented.

1. Introduction

It is commonplace for students to speak of an assessment task as being "easy" or "hard". Frequently, these judgements are at odds with the perception of the person setting the task. Academics will often express astonishment at students' inability to answer "easy" questions. In this paper, we investigate students' perception of the difficulty of a carefully chosen set of questions, with the aim of identifying the type of questions that students perceive as easy or difficult. This may enable us to modify our teaching practices and empower students to attack "difficult" tasks with more confidence.

In this paper we will consider the following: What types of questions do students perceive as difficult and what do they perceive as easy? Do their perceptions change after they have done the tasks? Are their perceptions based on familiarity with the type of question? Are their perceptions based on the conceptual difficulty of the question? Is the language of the question important? Do students with a non-English speaking background (NESB) or male/female students perceive questions differently? Is there a difference in performance in examinations on different types of questions?

In a previous paper (Smith *et al*, 1996), we developed a taxonomy to classify assessment tasks ordered by the nature of the activity required to successfully complete each task, rather than in terms of difficulty. The taxonomy, listed in Appendix C, has eight separate categories and we investigated the links between students' perceptions of difficulty and the categories of the taxonomy.

In relation to their perceptions of difficulty, we examined students' performance on two separate examination papers. One of these was designed to test factual knowledge, comprehension and use of routine procedures (Group A in the *MATH* taxonomy, Appendix C). This examination was of two hours duration and students had no aids. The second examination was designed to test higher-level skills (Groups B and C in the taxonomy). Students had three hours to complete the examination and could use one handwritten A4 sheet of notes and calculator. We believe that it is important to analyse the link between perception and success in assessment. Are their perceptions of difficulty born out by examination results? Do students avoid questions that they perceive as difficult?

Previous studies have considered students performance on statistics examinations in relation to the complexity of language in the question (Smith *et al*, 1994). This study of 186 students showed that there was no correlation between performance on examinations and the linguistic complexity of the question as measured by lexical density. This was also demonstrated with a study of 660 first year calculus students (Craig, T, 2001).

Craig's (2001) thesis considers calculus word problems and students perceptions of the difficulty of the problem. She looks at the variables of concrete versus abstract and the types of representation in the problem. The important variables for the perception of difficulty were familiarity of the problem, the context and whether there was a visual representation. Smith *et al* (1994) concluded that the conceptual difficulty of the mathematics was the important variable in the students' performance.

In the present study we consider a series of tasks requiring differing conceptual skills the students' perceptions of their difficulty. We examine the students' perceptions before and after completing the task to identify any changes that may have occurred. We look at students' performance on the end-of-semester examinations to see whether learning has occurred and to find which variables may cause significant differences in performance.

2. Method

Sample. Seventy students from a cohort of 90 in a first semester university calculus course were included in the study. The survey was voluntary and students were asked to sign an ethics approval form to use their data. Those that did not give approval were not included. There were 31 female students, 37 male students and 2 for whom this information was missing. Twenty-nine students spoke English at home, 39 spoke a language other than English and there were 2 missing data points. There were 13 students who had been in Australia less than 5 years.

Survey. The survey consisted of two parts. Students were asked to read a set of 8 questions (see Appendix A) and rank them in order of difficulty. The questions were representative of the eight categories of our taxonomy, but were presented in no particular order. The students also rated each question for skills required, level of difficulty, clarity and previous experience in answering those types of questions. Students were then asked to attempt the questions and re-rank them in order of difficulty. There was opportunity for open-ended comments. The survey items are listed in appendix B.

Questions. The questions were sample examination questions that the students had not seen but were related to the material they were studying in class. They were chosen as examples that would fit into the categories of the *MATH* taxonomy (Smith *at al*, 1996, Ball *et al*, 1998).

Examinations. As described in the introduction, we studied student performance on two different types of examination paper. The results were analysed for significant differences in student performance due to sex, language background and length of time in Australia.

3. Results

Examinations. Firstly the results of the examinations were analysed to determine if there were any differences between groups of students and to analyse whether learning had occurred. Data on sex, home language background and years in Australia were available. There was high correlation between the two examination results (0.67) and most students achieved satisfactory results. We can conclude that the majority of students reached the objectives of the subject. The only significant difference between groups was for the students who had recently arrived in Australia. Their results on examination 1 (routine skills) were significantly higher than for students, who had been in Australia longer (mean 64 for recent arrivals, mean 44 for others, $p = 0.012$). On the second examination paper (conceptual skills) there was no significant difference between the groups (mean 56 for recent arrivals, mean 59 for others). The students whose home language was not English also did better on the routine skills but this was not significant ($p = 0.062$). There was no significant difference in the sex and language interaction, as had been noted in earlier studies (Smith & Wood, 1998).

Rankings. The rankings before and after were analyzed and compared with the *MATH* taxonomy order (Table 1). There was no *a priori* reason for the *MATH* taxonomy rankings to reflect difficulty, since this was not the rationale for its development. Rather, it was designed to reflect conceptual complexity. There is considerable agreement between the taxonomy categories and the ranking given by the students. The 3 Group C categories were in the 4 questions perceived to be most difficult, the 3

Group A categories were in the 4 questions perceived to be easiest, while the 2 Group B categories were in the 4 questions perceived to be in the middle range of difficulty.

The pre- and post-rankings were compared using paired *t*-tests. The significant changes in rankings were:

Question B: harder ($p = 0.001$)

Question C: easier ($p = 0.012$)

Question G: easier ($p = 0.008$)

Taxonomy ranking	Pre ranking	Mean	Post ranking	Mean
C	F	2.59	C	2.64 **
F	C	3.37	F	2.87
A	H	4.26	G	3.77 **
D	A	4.54	H	4.67
H	G	4.56	A	4.71
G	D	4.96	D	4.71
B	B	5.06	B	6.03 **
E	E	6.67	E	6.57

Table 1: Rankings of difficulty of questions before doing the question and after. Significant change indicated by**.

Questions F and C were considered easy before and after doing the questions. Questions B and E were considered the most difficult before and after. The other questions were of a similar ranking before doing the questions. Of the middle group, only G changed significantly in the post ranking.

To investigate the reasons why students chose the rankings, we asked whether the questions were clearly worded, whether they understood the questions and whether they had see that style of question before. In each of these areas, there were significant differences in the responses over the 8 questions.

There were significant differences between the questions as to students' familiarity with the type of question. For example, question B (mean 2.3 on 5-point scale) was considered a familiar question but was ranked as difficult. Question A (mean 1.5 on 5-point scale) was considered very familiar but was not ranked as very easy. Students were familiar enough with the type of question to perceive that the presence of square root would increase the algorithmic complexity. Question D was ranked in the middle for perceived degree of difficulty but students had not seen this type of question before (mean 3.3). Question E was the most unfamiliar question (mean 3.7) and ranked the most difficult. B, C, F, G, H were assessed as having similar familiarity but were ranked very differently. The mean scores are presented below (Table 2).

The language is very clear (5-point scale, 1= very clear, 5= too hard to understand)

A	B	C	D	E	F	G	H
1.36	1.66	1.74	1.67	2.37	1.36	1.58	1.73

I understand the question (5-point scale, 1= I understand the question, 5= I do not understand the question)

A	B	C	D	E	F	G	H
1.32	1.94	1.78	1.54	2.41	1.30	1.58	1.69

Similar questions (5-point scale, 1= I have done similar questions before, 5= I have never done this type of question before)

A	B	C	D	E	F	G	H
1.46	2.35	2.66	3.30	3.74	2.25	2.70	2.35

Table 2: Mean scores for questionnaire analysis

Differences between students. The data were checked for differences between groups of students, in particular with regard to the variables sex, language background and years in Australia. There were no significant differences between male and female student for any rankings. There were significant differences between students who spoke languages other than English as their home language for question B (mean (NESB) = 5.19, mean (ESB) = 7.21, $p = 0.000$) and question D (mean (NESB) = 5.34, mean (ESB) = 3.83, $p = 0.001$). These differences persisted in the post rankings. When one looks at the question, it is not surprising that NESB students perceive question D to be difficult. It requires competence with English. It is not clear why there was such a significant difference between NESB and ESB students on question B but the simplicity of the language, that is very few words, may be the reason.

Open-ended comments. Students were invited to comment on their perceptions. They obviously enjoyed the task of ranking the questions and made some interesting comments. The comments are coded by the student number assigned as part of the anonymity provision of the ethics approval.

There were several students who generally underestimated the difficulty of the questions:

Reading questions may sometime seem easy but when you actually start to do them is when you start to see the difficulty. (12)

A question may look easy enough to do but applying all the information that you know to the question may be quite difficult. In all I underestimated what was asked of the question. (63)

We totally underestimated the hardness of the questions at first glance. Closer inspection of the question revealed the exact nature of the question. (72)

Yes my perceptions about the questions changed as a result of doing it because when I started to read or many other people started to read they got a misunderstanding of the question. Some questions are hard but first look very easy and vice versa. So when I actually sat down to do the question I found out it is harder than I expected it to be. (32)

Other students overestimated the difficulty of the questions:

When I first looked at the questions briefly they appeared quite hard, but on closer look and actually attempting them, they were actually relatively easy. (65) (68 very similar)

When I first looked at some of the questions they seemed really hard, but when I read over them and understood what they were asking, I found them less difficult than I originally thought. (36)

Some students found certain types of questions easier than they expected:

My perceptions have totally changed because the questions that dealt with definitions and explanations have tended to be easier than the questions where practice is necessary. (30)

My difficulty ranking has changed as a result of doing it. I thought that the questions that involved memorising facts or rules like E would be more difficult than other questions since it requires memory of facts/rules rather than logically proving. (42)

Some students realised that they needed to do more revision:

My perceptions about the questions have changed. They are not difficult to do if I had studied a bit more, or a whole lot more. (35)

Yes I read them and I understand what can I do but when I perform them I impact from problems like the rules, memories or calculation etc. (33)

My perceptions have changed because of lack of revision in the subject; I was unfamiliar with the types of questions asked. (46)

Comment from a NESB student who is articulating the difficulty with English that was demonstrated in the previous section.

The order of difficulty does not change much. The hardest question for me is still the theorem (E), i.e. language problem. All calculation is all right for me, except some question need to know more English. (66)

A couple of the students commented on question B. We think that they enjoyed solving it.

My perception didn't change much except that A was easier than I thought. B required a lot of thought – more than I expected. (73)

A was harder than I originally thought, B was impossible but I thought and worked it out, H got easier, D got harder. Basically first impressions don't really count. Only after close consideration can one judge the difficulty of a question. (60)

One student articulated the idea that a familiar question was easier. This was not demonstrated in the numerical data.

We may find some questions hard at the beginning because we think that we have never done that type of question before. (76)

4. Conclusions

We have found that students perceive questions to be difficult for a number of reasons. In general, questions requiring conceptual understanding are regarded as more difficult that those which require factual recall or the use of routine procedures. There was not a strong link between familiarity with the question type and ranking of difficulty. Students were sufficiently familiar with the some types of question to be able to perceive inherent difficulties, such as a complex differentiation.

We found that in five out of eight cases, students' perceptions of the difficulty did not change after they had done the task. In one case they found the question to be more difficult than expected and in two cases to be easier. It is not clear to us why students found Question B to be more difficult than expected. It may be that some of the complexities (such as the use of the intermediate value theorem)

were not immediately apparent. It is also significant that NESB students rated this question as easier than ESB students. This was the case both before and after attempting the question.

Not surprisingly, Question D, which required sophisticated language skills for its answer, was rated significantly more difficult by NESB students. Although the students understood what was being asked, they realised the need for language skills to answer it.

The close agreement between the *MATH* taxonomy and the ranking of difficulty given by the students is some evidence that the perceived difficulty is related to conceptual difficulty of the question.

The open-ended comments after completing the ranking showed that students found the exercise interesting and were surprised at the differences between their perceptions and the reality. Many commented that their ranking had not changed but that they had either underestimated or overestimated the difficulty of all the questions. Other students found that questions that dealt with definitions and theorems were easier than they expected. NESB students articulated their difficulties with answering questions, which required English skills.

REFERENCES

-Ball, G., Stephenson, B. Smith, G.H., Wood, L.N., Coupland, M. & Crawford, K., 1998, Creating a diversity of experiences for tertiary students, *Int. J. Math. Educ. Sci. Technol. 29*, 6, 827-841.

-Craig, T., 2001, *Factors affecting Perceptions of Difficulty in Calculus Word problems*. Unpublished Masters Thesis, University of Cape Town

-Smith, G.H., Wood, L.N., Coupland, M., Stephenson, B., Crawford, K. & Ball, G., 1996, Constructing mathematical examinations to assess a range of knowledge and skills *; Int. J. Math. Educ. Sci. Technol., 27*, 1, 65-77.

-Smith, N.F., Wood, L.N., Gillies, R.K. & Perrett, G., 1994, Analysis of students' performance in statistics, In Bell, G. Wright, B., Gleeson, N. & Geake, J. (eds) *Challenges in Mathematics Education: Constraints on Construction*. MERGA, 539-543.

-Smith, G.H. & Wood, L.N., 1998, Examination responses to changes in undergraduate mathematics assessment. *ICTM*, Wiley: Samos, 268-271.

Appendix A. Questions for ranking of difficulty.

A. Sketch the graph of $f(x)=\sqrt{\dfrac{x}{x+1}}$ showing the main features.

B. Show that $x^3 + cx + d = 0$ has only one root if $c \geq 0$.

C. What is the formula for the linear approximation to the function $f(x)$ at the point $x = a$?

D. Describe, in about 10 lines, the ideas of the mean value theorem. Imagine that you are describing the theorem to a student about to start university.

E. The mean value theorem is a powerful tool in calculus. List three consequences of the mean value theorem and show how the theorem is used in the proofs of these consequences.

F. Explain the differences between instantaneous velocity and average velocity.

G. Explain why the mean value theorem does not apply to the function $f(x)=|x+1|$ on the interval $[-3, 1]$.

H. Sketch a function $f(x)$ where $f(x)>0$, $f'(x)>0$ and $f''(x)>0$.

Appendix B. Questionnaire (data collected before students attempted the question)

1. I will need the following skills to answer this question. Feel free to circle more than one letter.
   (a) memorised facts and rules
   (b) the ability to justify what I am doing
   (c) practice in answering this type of question
   (d) the ability to describe what I am doing
   (e) the ability to apply my knowledge in a new situation
2. I would rate this question as:

Very easy	Easy	moderately hard	quite hard	impossible
1	2	3	4	5

3. The language is:

Very clear	clear	moderately hard	quite hard	Too hard to understand
1	2	3	4	5

4.

I understand the question				I do not understand the question
1	2	3	4	5

5.

I have done similar questions before				I have never done this type of question before
1	2	3	4	5

Appendix C. MATH Taxonomy (Smith *et al.* ,1996)

Group A	*Group B*	*Group C*
Factual knowledge (Question C)	Information transfer (Question D)	Justifying and interpreting (Question G)
Comprehension (Question F)	Applications in new situations (Question H)	Implication, conjectures and comparisons(Question B)
Routine use of procedures (Question A)		Evaluation (Question E)