Title

Does student confidence on multiple-choice question assessments provide useful information?

Journal

Medical Education, 47(6)

ISSN

0308-0110

Authors

Curtis, Donald A
Lind, Samuel L
Boscardin, Christy K
et al.

Publication Date

2013-06-01

DOI

10.1111/medu.12147

Peer reviewed

Does student confidence on multiple-choice question assessments provide useful information?

Donald A Curtis,[1] Samuel L Lind,[2] Christy K Boscardin[3] & Mark Dellinges[1]

CONTEXT Feedback from multiple-choice question (MCQ) assessments is typically limited to a percentage correct score, from which estimates of student competence are inferred. The students' confidence in their answers and the potential impact of incorrect answers on clinical care are seldom recorded. Our purpose was to evaluate student confidence in incorrect responses and to establish how confidence was influenced by the potential clinical impact of answers, question type and gender.

METHODS This was an exploratory, cross-sectional study conducted using a convenience sample of 104 Year 3 dental students completing 20 MCQs on implant dentistry. Students were asked to select the most correct response and to indicate their confidence in it for each question. Identifying both correctness and confidence allowed the designation of uninformed (incorrect and not confident) or misinformed (incorrect but confident) responses. In addition to recording correct/incorrect responses and student confidence, faculty staff designated incorrect responses as benign, inappropriate or potentially harmful if applied to clinical care. Question type was identified as factual or complex. Logistic regression was used to evaluate relationships between student confidence, and question type and gender.

RESULTS Students were misinformed more often than uninformed (22% versus 8%), and misinformed responses were more common with complex than factual questions ($p < 0.05$). Students were significantly more likely to be confident of correct than incorrect benign, incorrect inappropriate or incorrect harmful answers ($p < 0.001$), but, contrary to expectations, confidence did not decrease as answers became more harmful.

CONCLUSIONS Recording student confidence was helpful in identifying uninformed versus misinformed responses, which may allow for targeted remediation strategies. Making errors of calibration (confidence and accuracy) more visible may be relevant in feedback for professional development.

Medical Education 2013: ***47:*** *578–584*
doi: 10.1111/medu.12147

Discuss ideas arising from the article at www.meduedu.com 'discuss'

[1]Department of Preventive and Restorative Dental Sciences, School of Dentistry, University of California San Francisco, San Francisco, California, USA
[2]School of Economics and Business Administration, Saint Mary's College of California, Moraga, California, USA
[3]Office of Medical Education, School of Medicine, University of California San Francisco, San Francisco, California, USA

Correspondence: Donald A Curtis, Department of Preventive and Restorative Dental Sciences, D-3212, School of Dentistry, University of California San Francisco, 707 Parnassus, San Francisco, California 94143-0758, USA. Tel: 00 1 415 476 5827; E-mail: don.curtis@ucsf.edu

INTRODUCTION

Multiple-choice questions (MCQs) are helpful in providing a measure of what a student knows and, by subtraction, inferring what he or she does not know.[1,2] They are routinely used to assess knowledge of basic biomedical facts and, with additional effort, adapted to assess higher levels of learning, such as application and synthesis.[3,4] This has made MCQ assessments a versatile and reliable assessment strategy in medical education.[4–7]

There are continuing efforts to improve MCQs to provide more meaningful feedback to students, educators and testing agencies so that more valid inferences of competence can be determined and more learning can occur from the assessment process.[8–10] Asking students to consider their level of confidence in their answers has been recommended as a way to improve self-monitoring.[11,12] Similarly, asking students to reflect on the clinical impact of their decisions has been proposed as a way to make students more aware of the influence of their errors,[11–14] thereby improving the opportunity for learning.[9]

The primary advantage of recording both confidence levels and correctness on MCQ assessments is that it affords an ability to distinguish between students who are uninformed and those who are misinformed.[12,15] Students are considered to be uninformed when they select an incorrect answer and admit they are unsure. This combination of being incorrect and unsure is considered to provide a very appropriate 'teaching moment', in which the student is especially responsive to faculty feedback and to learning.[16] Similarly, a context in which a student has low confidence in a correct answer also represents an opportunity in which early feedback increases retention and improves metacognitive monitoring.[16,17] By contrast, students are misinformed when they select an incorrect answer, but state that they are sure or very sure of their response, which is qualitatively different from being uninformed. Strongly held incorrect beliefs are often resistant to change,[13,15–19] can interfere with student learning,[18,20] and may lead to inappropriate clinical decisions.[21]

The educational benefits to be derived from asking students to consider the level of confidence they have in their answers includes helping them identify limits of their knowledge,[15] reinforcement of the notion that guessing should be discouraged,[15] and the provision of an environment for active student self-monitoring.[9] Asking students to indicate their level of confidence in their answer to an MCQ can also cause them to spend more time on reviewing that item, especially if the error was unanticipated.[9] In this way, an examination can result in learning during an assessment, which is an example of test-enhanced learning.[8–10,22,23] Agrawal *et al.*[9] make the argument that testing identifies mistakes in accuracy, but also mistakes in self-monitoring, which is a valuable skill to develop and for which feedback is important. Feedback on self-monitoring may be helpful in metacognitive processing,[24] an issue especially problematic in underachieving students.[25] Confidence also influences how receptive a trainee is to feedback,[26] is generally lower on more complex process-type problems,[18] and may be influenced by gender.[27] The calibration of confidence to correctness in clinical decisions is a defining characteristic of an expert clinician and thus it is important to monitor and develop appropriate confidence in a trainee.[28]

Traditionally, assessments have been used to provide a quantitative measure of student knowledge, rather than a measure of student misunderstanding. We generally have no idea how sure students have been of their incorrect answers and therefore do not know when students are uninformed as opposed to misinformed. Additionally, by counting only correct responses, we assume all distractors are equally consequential, which is not likely. Some distractors may represent thinking that would result in a benign application to clinical outcomes, whereas others might result in inappropriate or potentially harmful clinical consequences. Understanding both the student's confidence in, and clinical impact of, incorrect answers on an MCQ examination would allow for a more complete appraisal of that trainee's professional development.

Evaluating both confidence in and the potential clinical impact of responses in MCQ assessments has not been previously attempted. Our purpose was to determine the percentage of incorrect answers in which the respondent had little confidence (uninformed) relative to that of incorrect answers in which the respondent was very confident (misinformed), and to establish whether student confidence was significantly impacted by the potential clinical consequences of answers, question complexity or gender.

METHODS

Participants

A total of 104 Year 3 dental students (58 male, 46 female) from the 2011 graduating class at the School of Dentistry, University of California San Francisco (UCSF), were included in this study. As part of their routine requirements, students completed a course on dental implants and an assessment that included 20 MCQs and a measure of how confident they were in their response to each question.

Assessment instrument

Clinical faculty staff chose 20 MCQs, which included 10 questions to evaluate knowledge of factual information and 10 to evaluate responses to complex concepts requiring interpretation or the analysis of a clinical scenario. Questions were selected from an existing bank of questions that had been used previously and for which reliability scores ranged from 0.65 to 0.80 and point biserial values were good. There were four possible responses for each question selected, including one most appropriate response and three incorrect distractors.

Clinical impact

Two faculty members independently designated the distractors as benign, inappropriate or harmful according to the following definitions: benign = results in an inconsequential or harmless patient outcome; inappropriate = either unsuitable or would delay the provision of appropriate patient care, and potentially harmful = results in direct and irreversible detrimental treatment of the patient. A kappa (κ) score, indicating the degree of agreement, was calculated for the two faculty members' concurrence of their designation of responses as benign, inappropriate or harmful for the 60 distractors across the 20 questions ($\kappa = 0.76$).

Student confidence

Students selected the answer they felt was most correct, but also indicated their level of confidence in their response to each item as: (i) very sure; (ii) sure; (iii) unsure, or (iv) very unsure. For statistical purposes, these categories were later dichotomised to 'confident' and 'not confident'. Students were told confidence levels would be queried as an important factor to consider, but were not told how measures of confidence would influence their grades.

Uninformed versus misinformed

By evaluating students' confidence in incorrect answers, responses were grouped as either uninformed or misinformed responses. Incorrect responses given by a student who lacked confidence were considered as uninformed responses; it was assumed that the student was either uncertain or guessing. Incorrect responses given by a student who was confident were considered as misinformed responses; the student was wrong, but felt strongly that the answer was correct.

Test results were collected using code numbers without identifiers to maintain confidentiality. The protocol for the study was reviewed and approved by the UCSF Committee of Human Research.

Data analysis

Descriptive statistical measures included the numbers and percentages of, respectively, correct, incorrect, uninformed and misinformed responses overall and by question complexity. A correlation between confidence and clinical impact was calculated for all 2080 responses using Spearman's rho (ρ).

Logistic regression (IBM SPSS Statistics for Windows, Version 20.0; IBM Corp., Armonk, NY, USA) was used, with question as the unit of analysis, resulting in 2080 observations (104 students $\times$ 20 questions). The dependent variable, confidence, was dichotomised into the categories of 'confident' and 'not confident' from the original polytomous ordinal student responses of 'very sure', 'sure', 'unsure' and 'very unsure'. The regression model analysed the relationships between student confidence and three non-parametric independent variables: potential clinical impact (benign, inappropriate, potentially harmful); question type (factual or complex), and gender. As question was the unit of measure and each student had more than one confidence score, confidence was not independent. To correct for this, we nested students ($n = 104$) by question ($n = 20$).

All three predictor variables were entered into the regression model as categorical covariates. Potential clinical impact measures of incorrect answers (benign, inappropriate and harmful) were

individually compared with the reference category, correct answers. Question type and gender were specified as binary indicator variables. The reference categories for question type and gender were 'complex questions' and 'male', respectively.

RESULTS

The mean score (correct) on the examination was 70% (range: 45–95%) (Table 1). Mean ± standard deviation (SD) student confidence, assessed on a 4-point scale, was 3.37 ± 0.80 overall, 3.53 ± 0.66 for correct answers and 3.08 ± 0.81 for incorrect answers. Mean ± SD student confidence was also calculated at each level of potential clinical impact as: benign incorrect answers, 2.51 ± 0.92; inappropriate incorrect answers, 3.02 ± 0.74, and harmful incorrect answers, 3.09 ± 0.94. When results were evaluated by total responses (n = 2080), 168 (8%) were identified as uninformed responses, in which students were incorrect and unsure, and 450 (22%) were identified as misinformed responses in which students were incorrect and confident (Table 1). Compared with simple questions, complex questions resulted in a lower percentage of correct responses (31% versus 39%) and a higher percentage of misinformed responses (13% versus 9%) (Table 1). When results were evaluated by student, the mean ± SD number of uninformed responses was 1.6 ± 1.6 (range: 0–7) and the mean ± SD number of misinformed responses was 4.3 ± 1.8 (range: 1–9). Correlations of confidence to clinical impact were moderate overall ($\rho = 0.31$, $p < 0.001$). The internal consistency for the 20 items within our study sample was 0.50 (Cronbach's alpha).

Regression analysis indicated students were significantly more likely to be confident in correct than incorrect benign (odds ratio [OR] 15.9, 95% confidence interval [CI] 8.1–31.2; $p < 0.001$), incorrect inappropriate (OR 3.6, 95% CI 2.7–5.0; $p < 0.001$), or incorrect harmful (OR 5.0, 95% CI 3.0–8.3; $p < 0.001$) answers. However, confidence did not decrease, as expected, with increasing levels of potential harm (Table 2). For example, students were more likely to be confident of correct than incorrect benign answers (OR 15.9), but less likely to be confident of correct than incorrect harmful answers (OR 5.0). This is also made apparent by the finding that the mean ± SD confidence level in incorrect benign answers (2.51 ± 0.92) was lower than that in incorrect harmful answers (3.09 ± 0.94). In other words, confidence did not decrease as answers became more dangerous (Table 2). Regression analysis also indicated that students were more likely to be confident of their responses to factual than complex questions (OR

Table 1 Percentages of total responses (n = 2080) identified as correct or incorrect, and as uninformed or misinformed, by question type and clinical impact

	Responses, %	Learning opportunity	Responses to simple questions, n *(%)*	Responses to complex questions, n *(%)*	Clinical impact		
Correct responses (n = *1462)*							
Correct and confident (knowledge)	64%	Minimal	752 (36%)	588 (28%)	No adverse clinical impact		
Correct and not confident (partial knowledge)	6%	Significant	54 (3%)	68 (3%)	No adverse clinical impact		
					Benign (n = *45)*	Inappropriate (n = *458)*	Harmful (n = *115)*
Incorrect responses (n = *618)*							
Incorrect and confident (misinformed)	22%	Refractory	179 (9%)	271 (13%)	22 (1%)	347 (17%)	81 (4%)
Incorrect and not confident (uninformed)	8%	Significant	55 (3%)	113 (5%)	23 (1%)	111 (5%)	34 (2%)

Table 2 Logistic regression results of student confidence in all responses (correct and incorrect), by potential clinical impact (benign, inappropriate, harmful), question type (factual or complex) and gender

	OR	95% CI	p-value
Incorrect and benign	15.9	8.1–31.2	< 0.001
Incorrect and inappropriate	3.6	2.7–5.0	< 0.001
Incorrect and harmful	5.0	3.0–8.3	< 0.001
Question type	1.4	1.0–1.8	< 0.05
Gender			NS

Odds ratios for potential clinical impact indicate the likelihood of a student being confident in a correct response compared with an incorrect response. Odds ratios by question type indicate the likelihood of a student being confident in a response to a factual question compared with a response to a complex scenario-type question.
OR = odds ratio; 95% CI = 95% confidence interval; NS = not significant

1.4, 95% CI 1.0–1.8; $p < 0.05$) and that gender was not a significant predictor of student confidence ($p > 0.05$) (Table 2).

DISCUSSION

Many of our findings were expected, but some were unanticipated. We were pleased that students were more confident of correct than incorrect answers ($p < 0.001$) and not surprised that students were more confident in their responses to factual than to complex questions ($p < 0.05$). Our finding that students reported greater confidence in correct than incorrect answers and that gender made no difference to this is consistent with recent literature.[9] However, we were surprised to find more misinformed (22%) than uninformed (8%) responses. Also unanticipated was the finding that student confidence did not decrease as the potential harm of answers increased.

A potential benefit of recording the correctness of and confidence in answers may be that the process provides detailed feedback for faculty staff, allowing a more targeted remediation. We are identifying gaps in knowledge when we differentiate correct from incorrect answers, but by recording confidence we are also differentiating the uninformed from the misinformed response. Uninformed students lack knowledge; they do not know something and are not confident. By contrast, students are misinformed when they select an incorrect answer, but state they are confident of their response. We found a mean ± SD of 4.3 ± 1.8 misinformed responses per student and a mean ± SD of 1.6 ± 1.6 uninformed responses per student. This is an important distinction because learning potential and remediation strategies for the two groups would differ.[13] Incorrect responses in which the student reports little confidence call for student remediation that simply requires the assimilation of additional knowledge. By contrast, students who select an incorrect answer but are confident of their response often strongly believe in incorrect information and may be resistant to change.[13,17] The first step in remediation would be to identify the students and specific topics that resulted in misinformed responses. A logical second step towards guiding students to recognise and learn from errors, such as strongly held incorrect beliefs, may involve scoring these responses lower, as suggested by Burton.[15] Additionally, feedback has been shown to improve student metacognitive monitoring, especially in students who give correct responses in which they have little confidence.[17] Lastly, awarding a higher potential score for complex scenario questions than for factual questions might be considered, given that more misinformed responses were recorded on to complex questions.

Student confidence was significantly influenced by question type. We noted more strongly held incorrect beliefs on complex questions than on factual questions (13% versus 9%), and our regression model showed students were more likely to be confident of factual than complex responses (OR 1.4, 95% CI 1.0–1.8; $p < 0.05$). Agrawal *et al.*[9] determined that confidence levels were not influenced by question type (factual versus vignette-based). Although guiding students to address complex questions in a systematic way has been shown to improve performance in some subject domains,[18] additional study is necessary to better understand the relationship between confidence and question type.

Students were more likely to be confident in correct than incorrect benign, incorrect inappropriate and incorrect harmful answers ($p < 0.001$), yet confidence did not decrease, as expected, as incorrect

responses became more dangerous (Table 2). For example, students were more likely to be confident of correct than incorrect benign answers (OR 15.9), but less so with harmful answers (OR 5.0). This may reflect our methodology because although we explained to students that confidence was an important factor to consider on each question, we did not emphasise that students should consider the potential clinical impact of incorrect answers. We assumed students would implicitly consider the 'dangerousness' of the response, but our findings suggest confidence was influenced more by correctness than dangerousness. It could be argued that recording potential clinical impact did not identify dangerous thinking because dangerous student responses may represent lack of knowledge rather than a deliberate act, as Slogoff and Hughes[29] have stated.

Assessments provide important benefits in helping to identify knowledge gaps, but also support the less appreciated and understood benefits of improving learning, as Roediger and Karpicke[1] have noted, and, importantly, in providing feedback on self-awareness during decision making, as Agrawal *et al.*[9] and Eva and Regehr[30] have noted. In the present study, we demonstrated that recording student confidence can help to identify uninformed versus misinformed responses. Asking students to consider their confidence and correctness concurrently may help students calibrate appropriate levels of confidence in what they do and do not know, thus developing an ability that represents an important characteristic in true experts.

LIMITATIONS OF THE STUDY AND QUESTIONS FOR FUTURE RESEARCH

This study was performed using a convenience sample, on one assessment with moderate reliability, at one institution, in one subject domain and was completed in one setting. Therefore, it must be considered a preliminary study. Additionally, this test format was time-consuming to develop and raised concerns in students provoked by the 'new' format. We are unsure why students gave more misinformed than uninformed responses, yet our finding would be consistent with those in Kruger and Dunning's work on being 'unskilled and unaware of it'.[25] Future studies will evaluate confidence over time and in different subject domains, with high and low achievers, and on assessments of varying difficulty. Despite this, our findings are consistent with those of other investigators on the correlation of confidence and correctness,[21] on the influence of question difficulty on confidence,[18] and in findings of high levels of confidence on incorrect answers[16] and no gender differences.[25]

CONCLUSIONS

Results from traditional MCQ assessments provide educators with details about what students know, but with less information about what they might do or what they are thinking. Asking students to state their confidence in their answers and recording the potential clinical consequences of incorrect answers change the traditional MCQ assessment from a 'forced choice' assessment of correctness to a measure of self-awareness, calibration and potential consequences. Students are asked to think about many more of the dimensions involved in clinical decision making, ideally early in their professional development. Our rationale in proposing a measure of potential clinical impact and confidence during formative assessments is that students in medical education should be thinking about the consequences and convictions of their decisions, even while they are in the classroom.

Students were more likely to be confident in correct answers than incorrect benign, incorrect inappropriate and incorrect harmful answers, yet confidence did not decrease, as expected, with the increasing dangerousness of answers.

The primary value of recording confidence on MCQ assessments may refer to its ability to make errors of calibration (confidence and accuracy) more visible to students early in their professional development.

A secondary purpose of recording confidence on MCQ assessments may be to provide feedback to faculty staff about the percentage of uninformed versus misinformed students in order that more targeted remediation strategies can be considered.

Contributors: DAC was responsible for the study design and wrote the first draft of the manuscript. SLL and CKB were responsible for data interpretation and analysis, and contributed to the writing of the manuscript. MD was responsible for the monitoring and collection of assessment data, and contributed to the writing of the manu-

script. All authors approved the final version of the manuscript submitted.

Acknowledgements: The authors would like to thank Patricia Cross, PhD, for valuable guidance on this manuscript.
Funding: none.
Conflicts of interest: none.
Ethical approval: this study was approved by the Committee of Human Research, University of California San Francisco.

REFERENCES

1 Roediger HL, Karpicke JD. Test-enhanced learning: taking memory tests improves long-term retention. *Psychol Sci* 2006;**17**:249–55.
2 Norcini JJ, McKinley DW. Assessment methods in medical education. *Teach Teach Educ* 2007;**23**:239–50.
3 Schuwirth LW, van der Vleuten CPM. Different written assessment methods: what can be said about their strengths and weaknesses? *Med Educ* 2004;**38**:974–9.
4 McCourbie P. Improving the fairness of multiple-choice questions: a literature review. *Med Teach* 2004;**26**:709–12.
5 Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non-functioning distractors in multiple-choice questions: a descriptive analysis. *BMC Med Educ* 2009;**9**:1–8.
6 Downing SM, Baranowski RA, Grosso LJ, Norcini JJ. Item type and cognitive ability measured: the validity evidence for multiple true-false items in medical specialty certification. *Appl Meas Educ* 1995;**8**:187–97.
7 Wass V, McGibbon D, van der Vleuten C. Composite undergraduate clinical examinations: how should the components be combined to maximise reliability? *Med Educ* 2001;**35**:326–30.
8 Eva KW. Diagnostic error in medical education: where wrongs can make rights. *Adv Health Sci Educ* 2009;**14**:71–81.
9 Agrawal S, Norman GR, Eva KW. Influences on medical students' self-regulated learning after test completion. *Med Educ* 2012;**46**:326–35.
10 Larsen DP, Butler AC, Roediger HL III. Test-enhanced learning in medical education. *Med Educ* 2008;**42**:959–66.
11 Graber ML. Educational strategies to reduce diagnostic error: can you teach this stuff? *Adv Health Sci Educ* 2009;**14** (Suppl):63–9.
12 Curtis DA, Lind SL, Dellinges M, Schroeder K. Identifying student misconceptions in biomedical course assessments in dental education. *J Dent Educ* 2012;**76**:1183–94.
13 Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. *Am J Med* 2008;**121** (Suppl):2–33.
14 Courtenay B, Weidemann C. The effects of a 'don't know' response on Palmore's Facts on Ageing Quizzes. *Gerontologist* 1985;**25**:177–81.
15 Burton RF. Misinformation, partial knowledge, and guessing in true/false tests. *Med Educ* 2002;**36**:805–11.
16 Wakabayshi T, Guskin K. The effect of an 'unsure' option on early childhood professionals' pre- and post-training knowledge assessments. *Am J Eval* 2010;**31**:486–98.
17 Butler AC, Karpicke JD, Roediger HL. Correcting a metacognitive error: feedback increases retention of low confidence correct responses. *J Exp Psychol Learn Mem Cogn* 2008;**34**:918–28.
18 Chi MT, Roscoe RD, Slotta JD, Roy M, Chase CC. Misconceived causal explanations for emergent processes. *Cogn Sci* 2012;**36**:1–61.
19 Sieck W, Arkes H. The recalcitrance of overconfidence and its contribution to decision aid neglect. *J Behav Decis Mak* 2005;**18**:29–53.
20 Kohnle A, Mclean S, Aliotta M. Towards a conceptual diagnostic survey in nuclear physics. *Eur J Phys* 2011;**32**:55–62.
21 Friedman CP, Gatti GG, Franz TM, Murphy GC, Wolf FM, Heckerling PS, Fine PL, Miller TM, Elstein AS. Do physicians know when their diagnoses are correct? Implications for decision support and error reduction. *J Gen Intern Med* 2005;**20**:334–9.
22 Larsen DP, Butler AC, Roediger HL III. Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Med Educ* 2009;**43**:1174–81.
23 Roediger HL, Marsh EJ. The positive and negative consequences of multiple-choice testing. *J Exp Psychol Learn Mem Cogn* 2005;**31**:1155–9.
24 Pallier G, Wilkinson R, Danthiir V, Kleitman S, Knezevic G, Stankov L, Roberts RD. The role of individual differences in the accuracy of confidence judgements. *J Gen Psychol* 2002;**3**:257–99.
25 Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognising one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol* 1999;**77**:1121–34.
26 Eva KW, Armson H, Holmboe E, Lockyer J, Loney E, Mann K, Sargeant J. Factors influencing responsiveness to feedback: on the interplay between fear, confidence and reasoning process. *Adv Health Sci Educ Theory Pract* 2012;**17**:15–26.
27 Koivula N, Hassmen P, Hunt DP. Performance on the Swedish Scholastic Aptitude Test: effects of self-assessment and gender. *Sex Roles* 2001;**44**:629–45.
28 Mamede S, Schmidt HG, Rikers R. Diagnostic errors and reflective practice in medicine. *J Eval Clin Pract* 2007;**13**:138–45.
29 Slogoff S, Hughes FP. Validity of scoring 'dangerous answers' on a written certification examination. *J Med Educ* 1987;**62**:625–31.
30 Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. *Acad Med* 2005;**80** (Suppl):46–54.

Received 11 June 2012; comments to authors 27 July 2012, 22 October 2012, accepted for publication 7 January 2013