Advances in Health Sciences Education

Medical student and junior doctors' tolerance of ambiguity: development of a new scale

--Manuscript Draft--

Manuscript Number:	AHSE-D-13-00406R3
Full Title:	Medical student and junior doctors' tolerance of ambiguity: development of a new scale
Article Type:	Original Paper
Keywords:	Ambiguity; epistemology; Medical Education; tolerance; uncertainty.
Corresponding Author:	Jason Hancock, MBChB Exeter Exeter, UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Exeter
Corresponding Author's Secondary Institution:	
First Author:	Jason Hancock, MBChB
First Author Secondary Information:	
Order of Authors:	Jason Hancock, MBChB
	Martin Roberts, BA (Hons), MSc
	Lynn Monrouxe, PhD
	Karen Mattick, PhD, Msc
Order of Authors Secondary Information:	
Abstract:	The practice of medicine involves inherent ambiguity, arising from limitations of knowledge, diagnostic problems, complexities of treatment and outcome and unpredictability of patient response. Research into doctors' tolerance of ambiguity is hampered by poor conceptual clarity and inadequate measurement scales. We aimed to create and pilot a measurement scale for tolerance of ambiguity in medical students and junior doctors that addresses the limitations of existing scales. After defining tolerance of ambiguity, scale items were generated by literature review and expert consultation. Feedback on the draft scale was sought and incorporated. 411 medical students and 75 Foundation doctors in Exeter, UK were asked to complete the scale. Psychometric analysis enabled further scale refinement and comparison of scale scores across subgroups. The pilot study achieved a 64% response rate. The final 29 item version of the Tolerance of Ambiguity in Medical Students and Doctors (TAMSAD) scale had good internal reliability (Cronbach's alpha 0.80). Tolerance of ambiguity was higher in Foundation Year 2 doctors than first, third and fourth year medical students (-5.23,P = 0.012; -5.98, P = 0.013; -4.62, P = 0.035, for each year group respectively). The TAMSAD scale offers a valid and reliable alternative to existing scales. Further work is required in different settings and in longitudinal studies but this study offers intriguing provisional insights.
Response to Reviewers:	Dear Sir, Thank you for your email providing further feedback on our second draft manuscript. We have now responded to your comments (see table below).We have included a new manuscript both with and without tracked changes. In the tracked changes version we have accepted all format changes and so tracked changes indicate changes to text and tables. We look forward to hearing from you in due course. Yours faithfully,

	Jason Hancock (corresponding author)

Tolerance ambiguity in medical students and doctors

Full title; Medical student and junior doctors' tolerance of ambiguity: development of a new scale

Jason Hancock[1], Martin Roberts[2], Lynn Monrouxe[3] & Karen Mattick[1]

1. University of Exeter Medical School, St Luke's Campus, Heavitree Rd, Exeter, EX1 2LU, UK

2. Collaboration for the Advancement of Medical Education Research and Assessment (CAMERA), Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK

3. School of Medicine, Cardiff University, Cardiff, UK

Correspondence address;

Dr. Jason Hancock;

University of Exeter, College House, St Lukes Campus, Heavitree Rd, Exeter, Devon EX1 2LU

j.hancock2@exeter.ac.uk, 01392 725955,

Abstract

The practice of medicine involves inherent ambiguity, arising from limitations of knowledge, diagnostic problems, complexities of treatment and outcome and unpredictability of patient response. Research into doctors' tolerance of ambiguity is hampered by poor conceptual clarity and inadequate measurement scales. We aimed to create and pilot a measurement scale for tolerance of ambiguity in medical students and junior doctors that addresses the limitations of existing scales. After defining tolerance of ambiguity, scale items were generated by literature review and expert consultation. Feedback on the draft scale was sought and incorporated. 411 medical students and 75 Foundation doctors in Exeter, UK were asked to complete the scale. Psychometric analysis enabled further scale refinement and comparison of scale scores across subgroups. The pilot study achieved a 64% response rate. The final 29 item version of the Tolerance of Ambiguity in Medical Students and Doctors (TAMSAD) scale had good internal reliability (Cronbach's alpha 0.80). Tolerance of ambiguity was higher in Foundation Year 2 doctors than first, third and fourth year medical students (-5.23,P = 0.012; -5.98, P = 0.013; -4.62, P = 0.035, for each year group respectively). The TAMSAD scale offers a valid and reliable alternative to existing scales. Further work is required in different settings and in longitudinal studies but this study offers intriguing provisional insights.

Key words;

Ambiguity, epistemology, medical education, tolerance, uncertainty,

Introduction

The practice of medicine involves inherent ambiguity and uncertainty, arising from limitations of knowledge, diagnostic problems, ambiguities of treatment and outcome, and unpredictability of patient response (Geller, Faden, & Levine, D, 1990). The ability of physicians to tolerate ambiguity is therefore of significant interest, with implications for doctors' mental health and wellbeing, staff retention in the medical profession, and specialty choice. For example, low tolerance of ambiguity has been linked with low patient and physician satisfaction, increased risk of physician burnout (Lim, 2003; Cooke, Doust, & Steele, 2013), more negative attitudes towards underserved populations (Wayne, Dellmore, Serna, Jerabek, Timm, & Kalishman, 2011), and personality traits such as dogmatism, conformity and rigidity (Budner, 1962; Furnham, & Ribchester, 1995). It has also been argued that evidence-based complex decision making, which requires the integration of individual patient perspectives and research evidence that may be incomplete, poor quality or conflicting, is only possible if the clinician is able to understand the limits of their own, and of scientific, knowledge and manage the associated uncertainty (Knight, & Mattick, 2006).

It is therefore understandable that undergraduate medical education has been encouraged to introduce educational strategies that will increase medical students' tolerance of ambiguity (Luther, & Crandall, 2011). However, the challenges are formidable. There is a fundamental lack of conceptual clarity around the term 'tolerance of ambiguity' and whether it can change over time in individuals or populations, or what strategies might enable it to change. There may also be unintended consequences associated with increasing all medical students' tolerance of ambiguity (Hancock, & Mattick, 2012). Crucially, the tools available to measure tolerance of ambiguity are crude, despite over 60 years of research, and this provides a particular barrier for the evaluation of educational strategies aimed at increasing learners' tolerance of ambiguity.

This paper aims to seek conceptual clarity around tolerance of ambiguity, to offer a measurement scale that can support the evaluation of educational strategies and to make use of a modern validity assessment framework to evaluate the validity of the scale for use in undergraduate medical students and foundation doctors. It also aims to make some tentative insights into whether the tolerance of ambiguity of populations of students changes during medical school.

Defining ambiguity and uncertainty

The Collins English dictionary defines ambiguity as “vagueness and uncertainty of meaning” and uncertainty as “not known, reliable or definite”. These definitions are conceptually similar and are often used interchangeably. However, a distinction has been made by some authors, including Greco & Roger (2002), who suggest that uncertainty is the response to an ambiguous situation, akin to the period of anticipation prior to the confrontation with a potentially harmful event. *Tolerance of ambiguity* has been defined as “the way an individual (or group) perceives and processes information about ambiguous situations or stimuli when confronted by an array of unfamiliar, complex or incongruent clues” (Furnham, & Ribchester, 1995). Indeed, an individual who is intolerant of ambiguity may experience stress when encountering an ambiguous situation, avoid ambiguous stimuli, seek clarity or act prematurely (Furnham, & Ribchester, 1995). *Intolerance of ambiguity* has been defined as the tendency to perceive or interpret ambiguous situations as actual or potential sources of psychological discomfort or threat (Norton, 1975). In these definitions, ambiguity can exist in a situation that is novel, unfamiliar or complex, or when the cues are contradictory. Other authors point out that tolerance of ambiguity might mean not only coping well in ambiguous situations but actively seeking out and thriving in them (Budner, 1962). This suggests the need for a potentially multidimensional construct of tolerance of ambiguity.

For this study, we used the Collins English dictionary definition of ambiguity and consider ambiguity to be the stimulus; and Greco et al.’s definition of uncertainty and consider uncertainty to be the response to an ambiguous situation. Therefore ambiguity and uncertainly are not fully synonymous, with tolerance of ambiguity being more wide ranging than tolerance of uncertainty, although they are closely related. In reality, it is likely that avoidance of uncertainty is correlated with intolerance of ambiguity (Furnham, & Ribchester, 1995). These definitions suggest that tolerance of ambiguity is closely aligned with personal epistemologies, although we are not aware of medical education literature that has explored this interface or overlap. The next section will explore the personal epistemological frameworks used by individuals when learning and making decisions, and the implications for this study.

Personal epistemology and tolerance of ambiguity - state versus trait?

Epistemology is a branch of philosophy that considers what it is to 'know': how we understand, integrate, justify, and apply knowledge. Early empirical data and theoretical models suggest that *personal epistemologies* change and develop over time, although there is less consensus about how this happens. Models of personal epistemology therefore describe development from a lay understanding of science, where science is considered to be based on certainties and 'truths', to an understanding that is more contextual and fluid (Knight, & Mattick, 2006; Norton, 1975). Research in this field has a long tradition starting with the works of Polanyi (1966) who tells us there is no such thing as objective factual science because all knowledge is understood through our own 'worldview', and Perry's developmental work on the nine stages of maturation (1968). However, in terms of the developmental perspective, early models tend to suggest linear development along an essentially unidimensional scale. More recent models challenge this view, suggesting that personal epistemology can have multiple dimensions, sometimes moving 'backwards' along the scale, and is likely to be topic / context specific. For example, in a recent study, first year medical students viewed anatomical knowledge as concrete and certain, whilst accepting that more ambiguity exists within the social sciences. In addition, students also progress to a more contextual and fluid understanding at a different rate for aspects of the same topic (Knight, & Mattick, 2006).

Hammer, and Elby (2002) agree that how an individual makes sense of a situation depends on the context of the situation but suggest individual variation in how the context activates the personal *epistemological resources* at their disposal. Some individuals may make use of a more static framework, based on beliefs surrounding the certainty and unchanging nature of science. Others make use of more dynamic framework, believing that scientific knowledge stems from evidence and changes and expands with everyday life. These frameworks are made up of different epistemological resources that we learn throughout our lives, including beliefs such as knowledge can be accumulated, and knowledge can be checked.

There are a number of implications of the work on personal epistemologies for our study of tolerance of ambiguity. First there is clear overlap of ideas and we would hypothesise that an individual with a more sophisticated personal epistemology would be more tolerant of ambiguity. Secondly, we conclude that we should be open to the possibility that tolerance of ambiguity will change over time, given appropriate environmental conditions and contextual exposure, rather than thinking of tolerance of ambiguity as a stable trait or personality variable as originally proposed by Budner in 1962. In other words, a change in learning context could encourage students to think differently about knowledge and evidence, and / or use different epistemological resources. This opens up exciting possibilities for educational strategies within medical school (Luther, & Crandall, 2011) perhaps involving early clinical and research exposure, and supported reflection. Finally, any attempt to measure a medical student's or doctor's tolerance of ambiguity would need to do so in clinical context, since the cues experienced in this context will activate the individual's epistemological resources and determine the level of ambiguity experienced by the individual practicing doctor.

Implications for measuring tolerance of ambiguity

From this analysis, we propose that a scale to measure tolerance of ambiguity amongst populations of medical students and junior doctors would need to: 1) contain items that are clinically contextualised; 2) have sufficient number and range of items to be sensitive to subtle changes; 3) be open to the possibility that tolerance of ambiguity is a multidimensional construct; and 4) demonstrate good validity evidence (Downing, 2003).

The most widely used scale to date is the original Budner scale (1962) or variations thereof. This scale has 16 items and good construct validity (Sidanius, 1988). However, it conceptualises tolerance of ambiguity as a single dimensional personality measure, the items are not clinically contextualised, and the internal reliability is poor (Cronbach's alpha 0.49 in the original report). Therefore this scale does not meet our four criteria above for a measurement tool for use with medical students. A 4 item modified Budner scale was used by Geller et al. (1990) which contains 3 items taken directly from the original Budner scale. Whilst two of the items were clinically contextualised and the Cronbach's alpha score did increase marginally (0.56), the other reservations still apply. In addition, the small number of items means the scale is unlikely to be sensitive to change. A more recent publication by Geller, Tambor, Chase, & Holtzmann, (1993) introduced a new scale which has an improved internal reliability (Cronbach's alpha 0.75) but is not clinically contextualized, contains only 7 items (reduced from 18 during the pilot study) and is unidimensional. The "Physicians reaction to uncertainty scale" (Gerrity, DeVellis, & Earp, 1990) comes closest to meeting our criteria, being clinically contextualised, containing 61 items and acting as a multidimensional measure. However the focus is very much on practicing physicians, is not validated for use in medical students and seeks to measure reaction to uncertainty rather than tolerance of ambiguity.

Therefore, to our knowledge, no scales exist that meet our requirements and can answer our question about change in tolerance of ambiguity during medical school.

Medical students' tolerance of ambiguity

Little empirical data exist about whether tolerance of ambiguity increases or decreases in medical students during undergraduate education and that which does exist are conflicting and potentially flawed due to the limitations identified in the measurement scales used. Budner (1962) suggests that tolerance of ambiguity may be higher in third year medical students than first years, although this was only found in one of the two medical schools studied and the differences were not statistical significant. In contrast, Geller et al. (1990) showed no difference between levels of ambiguity recorded in medical students (n=86) in their first, second and fourth year of study in a cross-sectional study in one medical school and concluded that selection may be more important than education and training in influencing tolerance of ambiguity amongst physicians. In postgraduate medical training, Deforge, and Sobal (1991) used the original Budner scale with 59 family practice residents in America and found higher tolerance of ambiguity in third year residents compared to first years. Their paper also claimed that medical students were more tolerant of ambiguity than the first year residents, however it is unclear if the 'slightly different' score referred to reached statistical significance. Recently, Geller hypothesised that those students entering medical school with a higher baseline tolerance of ambiguity may show an increase during medical school, while those less tolerant at baseline may show a reduction (Geller, 2013).

The tolerance of ambiguity of different sub-groups has also been explored, albeit with the same poor scales. Geller et al. (1990) reported that female medical students had a higher level of tolerance of ambiguity than male students, while Deforge, and Sobal (1991) found no gender differences in family practice residents. Medical students who were 'older' when they started medical school also had a higher tolerance of ambiguity, although the cut off for this is unclear (Geller et al., 1990).

Similarly, the association between tolerance of ambiguity and choice of medical specialty has been considered and this is perhaps the area of greatest consensus, although again we note the concerns about the data on which these conclusions are made. Budner (1962) reported that first to third year medical students who wished to pursue a 'less structured' career (e.g. psychiatry) were more tolerant of ambiguity; whereas those who wished to pursue the most structured careers (e.g. surgery, obstetrics and gynaecology) were more intolerant of ambiguity. Geller et al. (1990) confirmed that medical students wishing to pursue a career in psychiatry were more tolerant of ambiguity than those wishing to pursue a career in surgery. Furthermore Geller et al. (1993) reported that psychiatrists were more tolerant of ambiguity than obstetricians, paediatricians and family practitioners. Given that the entry criteria for core psychiatry training in the UK include the "capacity to deal with ambiguity & uncertainty in clinical life" then perhaps this is not surprising (Royal College Psychiatrists, 2013).

In summary based on the current literature, which is challenged by the methodological tools available, we can tentatively propose that tolerance of ambiguity may be associated with specialty choice and some demographic markers such as age and gender. It remains unclear whether tolerance of ambiguity changes during medical school but initial evidence suggests it may increase when working as a doctor, albeit a limited range of specialties have been explored.

Study aims and research questions

The aims of this study are twofold;

1. To design a measurement scale for tolerance of ambiguity in medical students and junior doctors that is clinically-contextualised but still relevant for first year medical students, that treats tolerance of ambiguity as a complex construct that may have multiple dimensions and be open to change, and that has a good internal reliability but has sufficient items that is likely to be sensitive to change. To use the results obtained to evaluate the validity of our scale in the population studied.
2. To offer some provisional insights into the associations between tolerance of ambiguity and stage in undergraduate / postgraduate training, demographics such as gender, entry status (e.g. prior degree) and prospective career choices.

The process of evaluating the validity of this scale involves developing a scientifically sound validity argument to support the intended interpretation of test scores for ambiguity and their relevance to the proposed use in the undergraduate medical student and foundation doctor population. The guidelines that we shall follow for this purpose are set out by the American Educational Research Association (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999) and have been applied directly to the medical education context by Downing (2003).

We hypothesise that tolerance for ambiguity will increase during medical education as students gain research skills and clinical experience, coupled with support for reflective practice, that allow them to develop new ways to activate and apply their epistemological resources; and that female participants who are older, have a prior degree on entering medical school, and/or wish to become psychiatrists will have a higher tolerance of ambiguity.

Methods

Item generation

First, a literature review was undertaken to identify the definitions, theories, empirical research and existing scales relevant to tolerance of ambiguity. Eight medical education colleagues at a medical school in South West England were then asked to send us scenarios highlighting examples of ambiguity in medicine and medical education. Using the themes identified from the literature and the scenarios as a guide, 49 draft items were generated (by KM and LM). Each item was placed into one of three subscales depending on what it was measuring; 'tolerance of ambiguity, seeks out and thrives in ambiguous situations', 'tolerance of ambiguity, but seeks to reduce ambiguity', or 'intolerance of ambiguity'. Individual scale items were developed, ensuring that they were clear, short, focused and had good face validity (Oppenheim, 2008). Four of these items were taken directly from the modified Budner scale used by Geller et al. in 1990 and a further 7 were taken from the original Budner scale (1962), although slightly modified to include the medical context. Therefore in total 10 items were taken from the original Bunder scale. All items were written as statements to which respondents were asked to score their agreement on a five-point Likert-type scale ranging from 'strongly disagree' (1) to 'strongly agree' (5), with the midpoint score being 'neutral' (3). This differs from the six point likert scales used previously by both Budner (1962) and Geller et al. (1990). The items focused on tolerance of ambiguity within a medical context, whilst ensuring the content was appropriate for first year medical students, but also relevant for junior doctors.

Item validation

The same eight colleagues reviewed the draft items. They were asked to give their opinion in the form of a numerical rating (1 to 5) on the quality of the item and asked to provide text comments on individual items in terms of (1) their relevance to the construct of tolerance of ambiguity; (2) their clarity; (3) their format; and (4) whether they would be understood by the range of target subjects, from first year undergraduate medical students to Foundation year doctors (junior doctors in their first two years after graduating from medical school). The experts were given the opportunity to reword or remove any items they felt inappropriate. If two experts suggested an item was removed then it was. Through this process, 9 items were removed, 14 more were reworded and 2 new items were written, resulting in an initial set of 42 items.

Ten medical students and Foundation year doctors reviewed the 42-item scale to ensure that the items were clear and understandable. In addition to scoring their agreement with each item, respondents were asked to provide free text comments and to identify poorly worded or difficult to answer items. Following this initial work (results not shown) one item was removed due to poor face validity and 2 more were reworded, resulting in a final 41-item TAMSAD questionnaire (Table 1).

Participants and recruitment

Following ethics approval, 411 undergraduate medical students across years 1-5 and 75 Foundation doctors, all based in Exeter, were approached and asked to complete the TAMSAD questionnaire. Participants were initially approached by e-mail with a link to a short YouTube video filmed by JH. Following this, the research team attended lectures and small group sessions and handed out paper versions of the questionnaire. Any medical student not seen was sent a copy of the TAMSAD by e-mail and asked to participate. All Foundation doctors were provided with electronic versions of the TAMSAD and asked to complete and return electronically. The questionnaire also included demographic questions, entry status (are you a graduate student?) and intentions regarding possible future careers in a range of specialties.

Statistical analysis

Data analysis was conducted in SPSS version 21. We calculated survey response rates within each year group. Negatively worded items were reverse scored prior to analysis. We calculated item response rates and Mahalanobis distances for each respondent and used these to identify, and remove where appropriate, potentially outlying respondents in the data set.

Responses to each item were examined for the distribution of responses, mean response score and standard deviation of response scores. If an item had a mean score below 2 or greater than 4, or if participants did not use the full five-point range of scale responses, then the standard deviation of the scores was examined. If the standard deviation was low, the items were considered for removal from further analysis.

The initial 41-item scale contained the 4 Geller et al. (1990) items and a further 7 items from the original 16 item Budner (1962) scale. We calculated Cronbach's alpha for each of these scales and compared with the values originally reported by those authors. Since we used only 10 of the 16 items from Budner's scale, if we include the 3 also used in the modified Geller scale, we used the Spearman-Brown formula to estimate Cronbach's alpha for the original scale length (Stanley, 1971). Due to the item differences and the incorporation of the items into a longer questionnaire these estimates are only proxies for the reliability of the Budner and Geller et al. scales in our sample.

Exploratory factor analyses using a variety of extraction and rotation procedures were used to investigate the possible existence of subscales within the overall scale. Following this, we developed a final TAMSAD scale by removing items which did not improve the value of Cronbach's alpha for the overall scale. We then removed items for which the adjusted item-total correlation with the remaining items was less than 0.20.

The final TAMSAD scale was used to conduct preliminary analyses on the levels of tolerance of ambiguity of medical students and doctors in the study sample. A TAMSAD score for each respondent was calculated as the mean item score (provided there were no more than two missing items) and linearly transformed from the original 1-5 scale to a 0-100 scale using the formula; New score = 25(Old score -1). The distribution of the scores across all respondents was examined for normality. Analysis of variance (ANOVA) was used to explore the possible influence on the

TAMSAD score of respondents' year group, gender, graduate entry status and interest in any of seven specialities (medicine, surgery, emergency medicine, general practice/ community medicine, psychiatry, paediatrics, radiology). We calculated effect sizes for independent predictors in relation to the magnitude of the standard deviation of the TAMSAD score (Cohen, 1988).

Results

Questionnaire analysis

Three of the 314 returned questionnaires had over half of the items unanswered and were excluded from the analysis. The Mahalanobis distance method identified eight potential outliers. One of these had employed an 'answer 1 or 5 strategy' and this questionnaire was also excluded from the analysis, giving an effective response rate of 310/486 (64%). Response rates varied by year group (Table 2).

Item 31 was removed from the TAMSAD scale because it had the lowest standard deviation of response scores (0.64) and on reflection the researchers felt it could be examining participant 'expertise' rather than their 'tolerance of ambiguity'. Since we wished to investigate the relationship between tolerance of ambiguity and specialty preferences we excluded items 5 and 15, which are clearly linked to such preferences.

The factor analysis indicated that the remaining 38 TAMSAD items could not be subdivided into a simple set of interpretable factors. Using principal factors extraction there were thirteen factors with eigenvalues greater than one but the scree plot suggested a five-factor solution accounting for 33% of the total variance. However the five factors enabled no simple interpretation (even after applying a Varimax rotation) and numerous items either had no factor loadings greater than 0.3 or loaded moderately on more than one factor (Appendix 1). Use of alternative extraction and rotation methods failed to find any simple solution. The initial Cronbach's alpha score was calculated at 0.75 and we interpreted this as suggesting that the TAMSAD questionnaire was acting as a unidimensional measure of tolerance of ambiguity.

We then looked to improve both the parsimony and reliability of the scale by reducing the total number of items. We found that removal of seven items (1, 2, 4, 12, 16, 34, & 40) increased the internal consistency of the scale to 0.80, while reducing the total number of questionnaire items to 31. Finally we removed two further items (14 and 37) which had adjusted item-total correlations less than 0.20. Note that item 11, which had an initial item-total correlation of 0.19 (Table 1), was retained at this stage as its item-total correlation with the remaining items was now above 0.20. This left the Cronbach's alpha unchanged at 0.80 indicating that the scale has a good internal consistency (Field, 2005) and could be interpreted as a unidimensional measure.

Three of these items had originally been intended to measure 'tolerance of ambiguity but seeks to reduce ambiguity' (4, 12, 16) rather than purely 'tolerance of ambiguity, seeks out and thrives in ambiguous situations'. Therefore it is not surprising that their removal from a scale that appears to be acting in a unidimensional way improves the internal consistency of the scale.

Table 1 shows items means and standard deviations for the original 41 items, indicates those which items came from the original Budner and Geller et al. scales, those which were retained in the final TAMSAD scale and their corrected item-total correlations with the scale.

We estimated the internal consistency reliability of the Geller et al. 4 item scale to be 0.31 and that of the full Budner 16 item scale to be 0.63.

Group differences

Participant scores on the 29-item TAMSAD scale ranged from 38.8 to 86.2 with a mean (SD) of 57.0 (8.8). Using the TAMSAD scale we found that significant differences in tolerance of ambiguity were associated with participants' year group but not with gender, graduate entry status or possible future career specialty interests (Table 3). Observed mean TAMSAD scores by year group are shown in Figure 1. First, third and fourth year medical students had significantly lower tolerance of ambiguity than FY2 doctors by 4.62 to 5.98 scale points. These are moderate effects (0.52 to 0.68 times the TAMSAD score SD) (Cohen, 1988). Tolerance of ambiguity in second and fifth year medical students and FY1 doctors was similar to that in FY2 doctors.

Participants expressing a preference for a possible career in surgery were, on average, 2.52 scale points lower in their tolerance of ambiguity than their peers, while those preferring paediatrics were 2.42 points higher but neither of these differences reached statistical significance at the 0.05 level. Prospective medics, GPs, emergency physicians, psychiatrists and radiologists had equivalent tolerance of ambiguity to their peers.

Discussion

This study aimed to design a scale to measure tolerance of ambiguity in medical students and junior doctors that address the limitations of existing scales, and to pilot it with the target population in one location. After several rounds of refinement, we arrived at a 29-item scale that we named the Tolerance of Ambiguity in Medical Students and Doctors (TAMSAD) scale.

We evaluated the validity of this scale using an established framework set out by the American Educational Research Association (AERA, APA, & NCME, 1999) and applied to the medical education context by Downing (2003). This framework states that when evaluating the validity of any assessment tool used in medical education five sources of evidence should be considered; content related validity evidence, the response process, the internal structure of the scale, the relationship to other variables and the consequences of using the assessment scale.

Content related validity evidence was provided through the provenance of the items, which were derived from an analysis of the education literature, from medical education theory and from existing tolerance of ambiguity scales. Since we did not want to assume that our scale would be acting as a unidimensional measure of ambiguity, items were initially separated into one of three subscales: 'tolerance of ambiguity, seeks out and thrives in ambiguous situations', 'tolerance of ambiguity, but seeks to reduce ambiguity', and 'intolerance of ambiguity'. Pilot work involved the input of academic staff working in medical education and from medical practitioners working in hospital and community settings. Unlike many previous scales, TAMSAD is context specific which allows it to assess an individual's tolerance of ambiguity in the medical setting. The pilot study achieved a 64% response rate across the five years of medical students and two years of Foundation doctors.

The response process was considered in the scale development stage as academics and clinicians were asked to remove or reword items that they felt inappropriate or difficult to understand. A pilot study was also completed during which 10 medical students and foundation doctors were asked to comment if they felt that items within the scale were difficult to understand or answer. Finally we ensured that data collected from the scales was accurately transcribed onto the statistical package by performing a thorough check of the data.

The internal structure of the TAMSAD scale was explored by measuring its internal reliability (Cronbach's alpha=0.80). We have interpreted this to mean that the scale is acting as a unidimensional measure of tolerance of ambiguity, which is supported by the improvement in internal consistency observed when we removed three of the four items initially created to measure 'tolerance of ambiguity, but seeks to reduce ambiguity'. The Cronbach's alpha associated with the TAMSAD scale in this population was higher than obtained by previous scales; 0.49 (Budner, 1962) and 0.56 (Geller et al., 1990). The Cronbach's alpha associated with the original Geller et al. questionnaire was lower than reported previously in our study population (0.31 in our study compared to 0.56). The expected Cronbach's alpha for the full Budner 16 item scale was 0.63, which was higher than Budner's reported value of 0.49, perhaps due to sample heterogeneity (i.e. if the participants in our sample were more varied in their tolerance of ambiguity than were those in Budner's sample then we would expect reliability to be higher).

The relationship of tolerance of ambiguity (as measured by TAMSAD) and other variables, such as stage of training, gender, graduate status and specialty choice, was sought. In our study, foundation year 2 doctors had a higher tolerance of ambiguity than first, third and fourth year medical students; other studies have reached various conclusions about association of tolerance of ambiguity and stage of training. Our study was a cross sectional survey, so conclusions about changes in tolerance of ambiguity over time cannot be made. Wayne et al. (2011) demonstrated that those students with a higher tolerance of ambiguity at the start of medical school showed a smaller deterioration in their attitudes towards underserved populations and other studies have shown that intolerance of ambiguity is associated with distress (Benbassat, Baumal, Chan, &Nirel, 2011) and reduced levels of work satisfaction (Bovier, & Perneger, 2007). In response, Luther et al. (2011)

recommended that medical schools do more to increase their students' tolerance of ambiguity but a responding commentary cautioned about the possible unintended consequences of doing this, for example resulting in an undersupply of surgical trainees. In reality it is likely that there are multiple factors underpinning any increase in tolerance of ambiguity during medical school, including the increasing maturity of students. Our study showed no significant association between tolerance of ambiguity in prospective surgeons or psychiatrists compared with their peers. Previous research has suggested an association between tolerance of ambiguity and medical career intention (Budner, 1962; Geller, Tambor, Chase, & Holtzmann, 1993).

We would argue that the consequences of completing this questionnaire are minimal. Our research suggests that completing the modified twenty-nine questionnaire will take 5–10 minutes and is unlikely to have a negative impact on participants.

The most important contribution of this study is to provide a valid tool for the research community to apply in subsequent studies. The provisional findings from piloting the scale are broadly supportive of previous research by Budner (1962) and Geller et al. (1990).

As with all studies, our work has a number of methodological strengths and challenges. The strengths of this study include the way in which we carefully defined the constructs of ambiguity and uncertainty, thus ensuring a rigorous process in the development of the scale. Additionally, there were multiple rounds of scale refinement based on target group feedback and on psychometric analysis. The pilot study achieved a large sample size and good response rate. In terms of challenges, data collection only took place in one site, and the study was cross sectional in design rather than longitudinal. These both serve to limit the conclusions that can be drawn in relation to change over time.

Future research

Further work to provide additional evidence for the validity of the TAMSAD scale is now required. One aspect that has not been explored in depth is the cultural sensitivity of the scale. The current study has used the scale in one location in South West England; future work could use the scale in different countries and with more culturally heterogonous populations. The fact that the scale seemed to be acting as a unidimensional measure was unexpected, given the theoretical complexity of the construct of tolerance of ambiguity. It would therefore be helpful to repeat the exploratory factor analysis process with other, larger populations of students to verify this observation. Further qualitative research could also be useful to explore the different aspects of the construct of tolerance of ambiguity in different settings and with different populations.

It will also be interesting to explore associations between tolerance of ambiguity (as measured by TAMSAD) with other variables such as attitudinal markers (e.g. cynicism), observed behaviours (e.g. medical professionalism), cognitive states (e.g. intellectual development) and aspiration (e.g. specialty choice). Specific variables of interest could include attitudes towards underserved populations (Wayne, 2011) and intellectual maturation (Perry, 1968).

Other work should develop the TAMSAD for use in an online forum. We were able to ensure good response process validity as the scale was only completed on a small scale, mainly on paper. If the scale is to be used on a larger population, then the use of an electronic scale would help to ensure that there are minimal errors when processing data.

Finally, following further validation, this scale could be used in a number of different ways to shed light onto how tolerance for ambiguity might change during medical undergraduate and early postgraduate training. Whilst our cross sectional survey design has provided interesting data, future research involving longitudinal methodologies would enable us to track students through medical school and into early clinical practice, therefore providing insights into the pattern of growth or decline in tolerance of ambiguity during medical school and junior doctor training. Such studies might enable us to draw conclusions about how levels of tolerance of ambiguity vary across different medical schools using different curricula.

Conclusion

The TAMSAD scale developed through this study offers a more valid and reliable alternative to existing scales for medical students and junior doctors. Further work is now required to continue the process of evaluating the validity of this scale in the undergraduate and foundation doctor population. This will be possible through conducting longitudinal studies to explore changes in tolerance of ambiguity, both over time and as a result of educational interventions. Meanwhile this study offers intriguing provisional insights that warrant further investigation.

Acknowledgements

We would like to thank everybody who contributed towards each stage of this study. Our eight medical education colleagues who contributed towards initial item creation, the ten medical students and foundation doctors in Exeter who took part in the initial study pilot and the medical students and foundation doctors in Exeter who completed the TAMSAD questionnaire.

References;

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association, 1999.

Benbassat, J. Baumal, R. Chan, S. &Nirel, N. (2011). Sources of distress during medical training and clinical practice: Suggestions for reducing their impact. *Medical Teacher*, 33(6), 486-490.

Bovier, P. &Perneger, T. (2007). Stress from uncertainty from graduation to retirement – a population-based study of Swiss physicians. *Journal of General Internal Medicine*, 22(5), 632-638.

Budner, S. (1962). Intolerance of ambiguity as a personality variable. *Journal of Personality*, 30(1), 29-50.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences.* (Lawrence Erlbaum Associates).

Cooke, G. Doust, J. & Steele, M. (2013). A survey of resilience, burnout, and tolerance of uncertainty in Australian general practice registrars. *BMC Medical Education*, 13, 2.

Downing, S. (2003). Validity: on the meaningful interpretation of assessment data. *Medical Education*, 37, 830-837.

Field A. (2005). *Discovering statistics using SPSS*, Second edition. (Sage publications).

Furnham, A. & Ribchester, T. (1995). Tolerance of ambiguity: A review of the concept, its measurement and applications. *Current Psychology*, 14(3), 179.

Geller, G. Faden, R. & Levine, D. (1990). Tolerance for ambiguity among medical students: implications for their selection, training and practice. *Social Science & Medicine*, 31(5), 619-624.

Geller, G. Tambor, E. Chase, G. &Holtzmann, N. (1993). Measuring physicians' tolerance for ambiguity and its relationship to their reported practices regarding genetic testing. *Medical Care*, 31(11), 989-1001.

Geller, G. (2013). Tolerance for ambiguity: An ethics-based criterion for Medical Student selection, *Academic Medicine*, 88(5), 581-584.

Gerrity, MS. DeVellis, RF. Earp, JA. (1990). Physicians' reactions to uncertainty in patient care. A new measure and new insights. *Medical Care*, 28(8), 724-736.

Greco, V. & Roger, D. (2002) Uncertainty, stress and health. *Personality and Individual Differences*, 34(6), 1057-1068.

Hammer, D. & Elby, A. (2002). On the form of a Personal Epistemology. (In Hofer, B. &Pintrich, P. (Eds.), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing* (pp169-190). Mahwah, NJ: Erlbaum.).

Hancock, J. & Mattick, K. (2012). Increasing Students' Tolerance of Ambiguity: The Need for Caution, Letters to the Editor. *Academic Medicine*, 87(7), 834.

Knight, L. & Mattick, K. (2006). When I first came here, I thought medicine was black and white: making sense of medical students' ways of knowing. *Social Science & Medicine*, 63(4), 1084-96.

Lim, M. (2003, April). Who Is Being Difficult? Addressing the Determinants of Difficult Patient-Physician Relationships.Virtual Mentor, 5, 4. Retrieved 13 June 2013, from http://virtualmentor.ama-assn.org/2003/04/jdsc2-0304.html

Luther, VP. & Crandall, SJ. (2011). Commentary: ambiguity and uncertainty: neglected elements of medical education curricula? *Academic Medicine*, 86(7), 799-800.

Norton, R. (1975). Measurement of ambiguity tolerance. *Journal of Personality Assessment*, 39(6), 607–619.

Oppenheim, AN. (2008). *Questionnaire design, interviewing and attitude measurement*. (London: Continuum).

Perry, W. (1968). *Forms of Intellectual and Ethical Development in the College Years – A Scheme*. (New York: Holt, Rinehart and Winston).

Polanyi, M. (1966). *The Tacit Dimension*. (London: Routledge).

Royal College Psychiatrists CT1 personal specification. Retrieved 24 March 2013 from;

http://www.rcpsych.ac.uk/pdf/CT1%20Psych-person%20spec-August%202013-approved.pdf

Sidanius, J. (1988). Intolerance of ambiguity, conservatism and racism: Whose fantasy, whose reality? A reply to Ray. *Political Psychology*, 9, 309-316.

Sobal, J. &,Deforce, B. (1992). Reliability of Budner's intolerance of ambiguity scale in medical students. *Psychological Reports*, 71, 15-18.

Stanley, J. (1971). *Educational Measurement*. Second edition. (Washington, DC: American Council on Education).

TAMSAD recruitment video. Retrieved 13 June 2013 from;

http://www.youtube.com/watch?v=uQEANkFzfq4&feature=youtu.be

Wayne, S. Dellmore, D. Serna, L. Jerabek, R. Timm, C. & Kalishman, S. (2011). The association between intolerance of ambiguity and decline in medical students attitudes toward the underserved. *Academic Medicine*, 86(7), 877-882.

Figure 1. Observed mean tolerance of ambiguity score (with 95% confidence interval) by year group. FY = Foundation year.

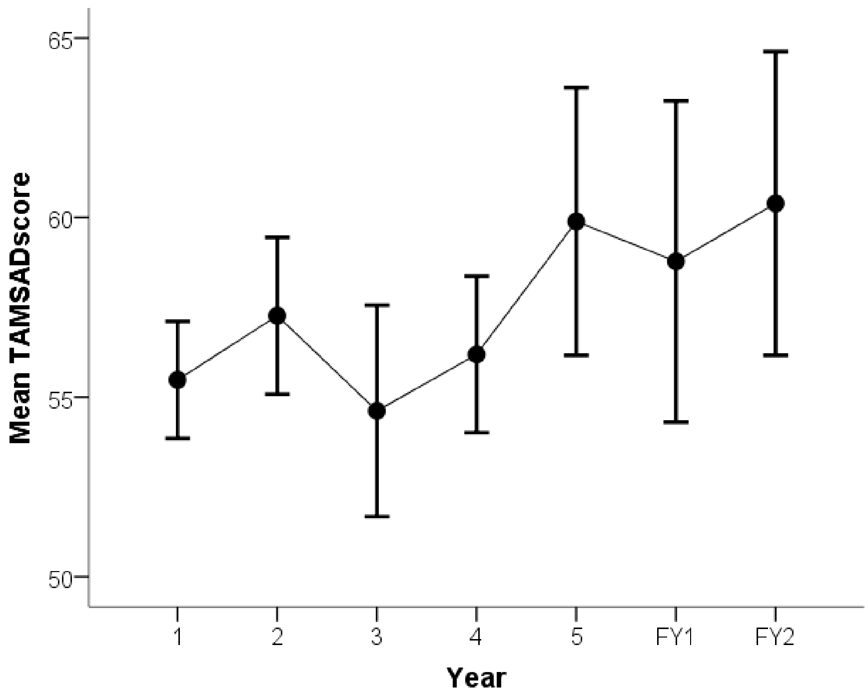


Table1; The TAMSAD scale.

	Item	Mean (likert score)	SD (likert score)	Was the item included in the final 30 item scale?	Corrected item-total correlation
1	I am comfortable to acknowledge that I'll never know everything about medicine.	4.28	.97	No	.13
2	Even when there is conflicting information, I prefer to make a decision and move on.*	3.27	.94	No	.01
3	I would enjoy tailoring treatments to individual patient problems.	4.11	.69	Yes	.30
4	I think it is important to attribute a percentage likelihood to a diagnosis or a specific patient outcome.	3.45	.91	No	-.05
5	As a doctor I would prefer the clear and definite work of someone like a surgeon to the uncertainties of a psychiatrist.*,$^{\alpha}$	3.14	1.24	No	.30
6	I have a lot of respect for consultants who always come up with a definite answer.*	3.65	.91	Yes	.40
7	I would be comfortable if a clinical teacher set me a vague assignment or task.$^{\pi}$	2.46	1.01	Yes	.30
8	A good clinical teacher is one who challenges your way of looking at clinical problems.$^{\pi}$	4.31	.67	Yes	.27
9	What we are used to is always preferable to what is unfamiliar.*,$^{\alpha,\pi}$	3.16	1.13	Yes	.24
10	I feel uncomfortable when people claim that something is 'absolutely certain' in medicine.	3.51	1.01	Yes	.24
11	A doctor who leads an even, regular work life	2.91	1.01	Yes	.19

	with few surprises, really has a lot to be grateful for.*,π				
12	I enjoy reducing the complexity of medical information to something more tangible.	3.95	.80	No	-.08
13	I think in medicine it is important to know exactly what you are talking about at all times.*	3.22	1.03	Yes	.32
14	'I don't know' are really important words in medicine.	4.19	.81	No	.18
15	I would prefer to work in a medical specialty where patients normally get better after treatment.*	3.92	.85	No	.24
16	I enjoy reducing detailed scientific problems to their core concepts.	3.91	.84	No	-.02
17	I feel comfortable that in medicine there is often no right or wrong answer.	3.64	.97	Yes	.41
18	A patient with multiple diseases would make a doctor's job more interesting.	3.51	.87	Yes	.30
19	I am uncomfortable that a lack of medical knowledge about some diseases means we can't help some patients.*	3.49	.95	Yes	.24
20	The unpredictability of a patient's response to medication would bring welcome complexity to a doctor's role.	2.84	.88	Yes	.24
21	It is important to appear knowledgeable to patients at all times.*	3.37	.99	Yes	.34
22	Being confronted with contradictory evidence in clinical practice makes me feel uncomfortable.*	3.03	.93	Yes	.41

23	I like the mystery that there are some things in medicine we'll never know.	3.20	1.13	Yes	.34
24	Variation between individual patients is a frustrating aspect of medicine.*	2.20	.95	Yes	.43
25	I find it frustrating when I can't find the answer to a clinical question.*	3.82	.80	Yes	.40
26	I am apprehensive when faced with a new clinical situation or problem.*,$^{\pi}$	3.10	1.00	Yes	.31
27	I feel uncomfortable knowing that many of our most important clinical decisions are based upon insufficient information.*	2.95	.86	Yes	.32
28	No matter how complicated the situation, a good doctor will be able to arrive at a yes or no answer.*,$^{\pi}$	2.30	.93	Yes	.28
29	I feel uncomfortable when textbooks or experts are factually incorrect.*	3.76	.92	Yes	.24
30	There is really no such thing as a clinical problem that can't be solved.*,$^{\pi}$	2.22	.86	Yes	.30
31	It's an exciting feeling when you listen to a patient tell you their symptoms and you just know what disease it is.*	1.6	.64	No	.06
32	I like the challenge of being thrown in the deep end with different medical situations.	3.52	.88	Yes	.33
33	It is more interesting to tackle a complicated clinical problem that to solve a simple one.$^{\pi}$	3.58	.96	Yes	.24
34	In medicine as in other professions, it is possible to get more done by tackling small, simple	3.33	.85	No	.00

	problems rather than large and complicated ones.*[α,π]				
35	I enjoy the process of working with a complex clinical problem and making it more manageable.	3.88	.67	Yes	.20
36	A good job is one where what is to be done and how it is to be done are always clear.*[α,π]	2.98	1.04	Yes	.40
37	Medicine has a lot of grey areas because we haven't found the answers yet.*	4.11	.83	No	.18
38	To me, medicine is black and white.*	1.58	.73	Yes	.28
39	The beauty of medicine is that it's always evolving and changing.	4.27	.70	Yes	.35
40	I enjoy working out which opinion is right in situations where many different opinions are expressed.*	3.69	.77	No	-.19
41	I would be comfortable to acknowledge the limits of my medical knowledge to patients.	4.01	.86	Yes	.27

** Negatively worded items that were reverse scored prior to analysis.*

[π] *Taken from the original Budner (1962) scale.*

[α] *Taken from the Geller et al. (1990) scale.*

Table 2, Response rates by year group.

Stage / year of training	Response rate	%
Undergraduate		
1	74/ 110	67%
2	72/112	64%
3	34/72	47%
4	52/78	67%
5	30/39	77%
Foundation training		
F1	22/39	56%
F2	26/36	72%
All	310/486	64%

Table 3. ANOVA results. Dependent variable: tolerance of ambiguity score.

Factor	Difference in TAMSAD score	P-value	LCL [a]	UCL [b]	Adjusted mean TAMSAD score [c]
Gender					
Female	1.03	0.329	-1.04	3.09	59.63
Male	Reference category		-	-	58.60
Entry status					
Graduate entry	2.27	0.168	-0.96	5.49	60.25
Non-graduate entry	Reference category		-	-	57.98
Year group		0.044 [d]			
Year 1	-5.23	0.012	-9.33	-1.14	57.11
Year 2	-3.87	0.065	-7.98	0.24	58.47
Year 3	-5.98	0.013	-10.67	-1.28	56.36
Year 4	-4.62	0.035	-8.92	-0.33	57.72
Year 5	-0.08	0.973	-4.77	4.61	62.26
FY1 [e]	-2.80	0.277	-7.87	2.26	59.54
FY2 [e]	Reference category		-	-	62.34
Specialty preferences [f]					
Medicine	0.92	0.387	-1.17	3.02	59.57
Surgery	-2.52	0.055	-5.09	0.05	57.85
Emergency	1.46	0.214	-0.85	3.77	59.84
GP / community	-1.18	0.287	-3.35	0.99	58.52
Psychiatry	2.22	0.233	-1.44	5.89	60.23
Paediatrics	2.42	0.062	-0.12	4.96	60.32
Radiology	-0.96	0.793	-8.19	6.27	58.63

a. LCL = Lower 95% confidence limit for the difference in score
b. UCL = Upper 95% confidence limit for the difference in score
c. Estimated marginal means, adjusted to account for between-group differences in the other variables in the model
d. P-value for F test of year group as a factor
e. FY = Foundation year
f. Specialty preferences were not exclusive. The reference category is ‘No interest in the specialty’ in all cases. Adjusted means for respondents not interested in each specialty can be obtained by subtracting the difference given in column 2 from the adjusted mean given in column 6.

Appendix 1. Factor analysis.

Rotated Factor Matrix

(Loadings <0.1 suppressed)

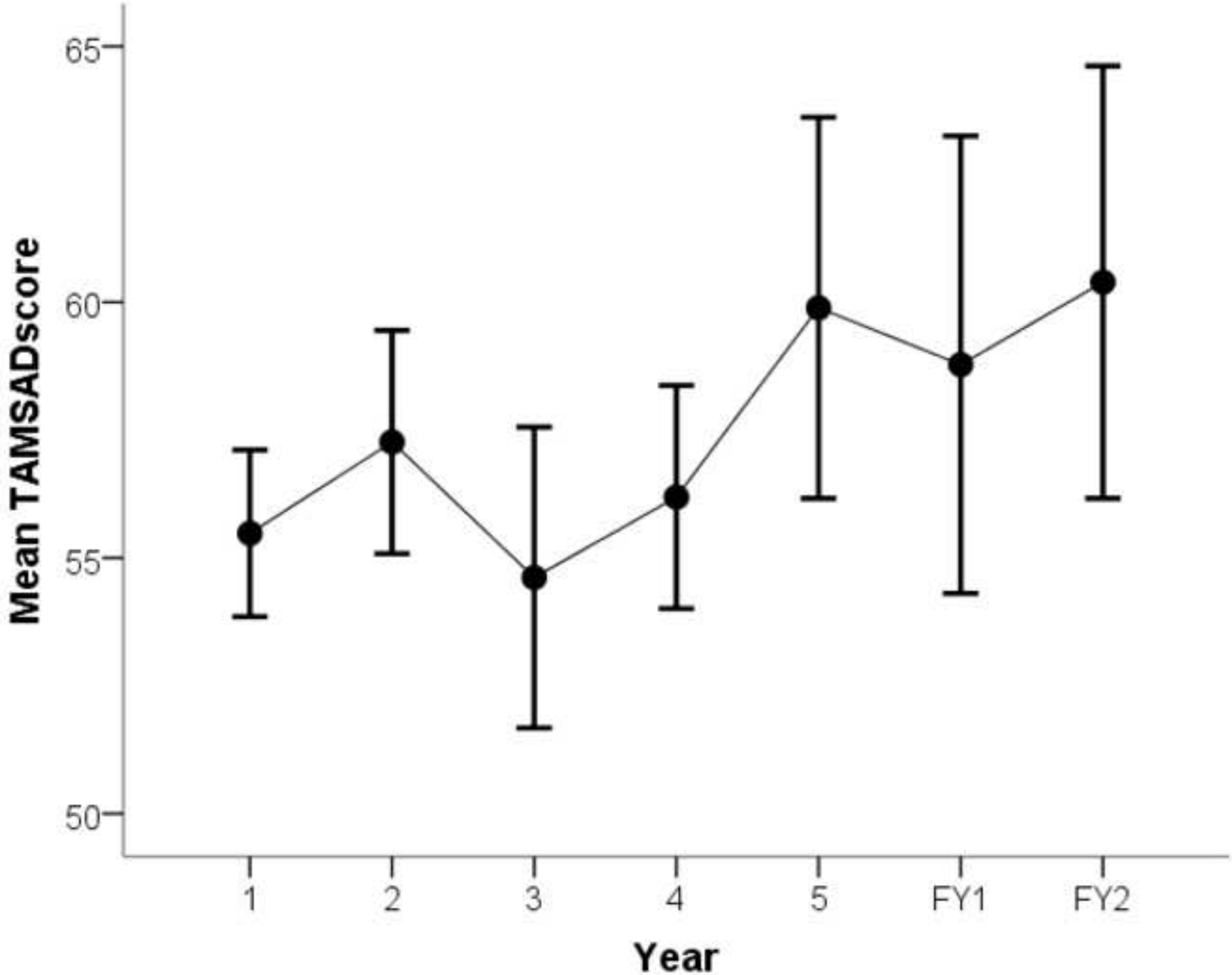
	Factor				
	1	2	3	4	5
Q38	.555	.107			
Q28	.518			.115	.170
Q21	.435		.204		
Q06	.433		.155	.101	.242
Q30	.429				
Q14	.391			-.250	
Q36	.373	.205	.140		.148
Q13	.371		.237		.142
Q41	.338		.160		-.150
Q03	.298	.261			-.151
Q01	.253		.231		-.230
Q10	.206	.106	.105	.102	
Q33		.629			
Q08	.126	.510			-.160
Q18		.506	.116		.121
Q35		.437			-.176
Q39	.144	.423	.153	.213	
Q24	.295	.369	.115	.145	.146
Q25			.548	.137	.181
Q23	.111		.499		

Q17	.256	.100	.424	.122	-.123
Q19			.356		
Q20		.200	.315		.180
Q29			.292		.184
Q32		.379	.166	.522	
Q11				.465	.232
Q26		.237	.163	.445	
Q22	.204		.284	.436	
Q40		-.313		-.327	.122
Q07			.249	.315	
Q27	.159		.289	.315	
Q02				-.215	.477
Q12					-.403
Q37			.101		.294
Q16		.199	-.109		-.289
Q34					.268
Q09	.117	.163		.157	.236
Q04					-.153

Extraction Method: Principal Axis Factoring.

Rotation Method: Varimax with Kaiser Normalization.

Figure
Click here to download high resolution image



15/4/14

Dear Sir,

Thank you for your email providing further feedback on our second draft manuscript. We have now responded to your comments (see table below).We have included a new manuscript both with and without tracked changes. In the tracked changes version we have accepted all format changes and so tracked changes indicate changes to text and tables.

We look forward to hearing from you in due course.

Yours faithfully,

Jason Hancock (corresponding author)

COMMENTS FOR THE AUTHOR:	Authors' response to comments
Editor	
I noted the addition of the item-total correlations to Table 1 and would suggest your remove the column labelled as corrected item-total correlation (final scale) and keep the one labelled as corrected item-total correlation (all 41 items). You could remove the "all 41 items" bit from the column heading because would be redundant with the other item-total correlation column removed.	The logic behind including the "corrected item-total correlation (final scale)" column was that it helped explain why item 11 was retained in the final scale. However, we have now explained this in an alternative manner (see below) and have removed this column. We have also removed the ''all 41 items'' bit from the remaining column heading as advised.
You commented that your criteria was an item-total correlation that was less than .2 but that item 11 was kept in although it had a correlation of .19. I assume that someone reviewed the items before removing and if so, then this should be no problem. To handle it, I would suggest the following edit. - Page 12 3rd paragraph. Change the wording of the second sentence to " Following this, we developed a final TAMSAD scale by first reviewing any items which did not improve the value of Cronbach's Alpha for the overall scale, followed by any items for which the adjusted item-total correlation was less than 0.20 If, upon review, the item was deemed as being ambiguous or otherwise inappropriate, it was removed from the scale" If on the other hand, removing items was a blanket rule that was applied without actually reviewing the items, then I would suggest leaving the wording on Page 12, 3rd paragraph as is but alter the wording on page 14, first paragraph to the following. "Finally, we removed two further items .... Note that Question 11 had an item-total correlation of .19. This item was kept in the scale because we viewed it as being important enough to warrant inclusion. This left the Cronbach's alpha......"	Neither of the reviewer's suggested interpretations of why item 11 remained in the final scale are quite in line with what we actually did. We have now amended the text of the Methods to clarify the fact that we applied the "item-total correlation < 0.20" criterion after having recalculated the item-total correlations once we had removed the items that were attenuating the value of Cronbach's alpha. The relevant wording now reads: "*Following this, we developed a final TAMSAD scale by removing items which did not improve the value of Cronbach's alpha for the overall scale. We then removed items for which the adjusted item-total correlation with the remaining items was less than 0.20.*" We have also re-emphasised this in the Results section, explaining why item 11 was retained using the following wording:: "*Note that item 11, which had an initial item-total correlation of 0.19 (Table 1), was retained at this stage as its item-total correlation with the remaining items was now above 0.20.*"
Thank your for adding the means to Table 3 but unfortunately they are not clear to me. The table labels them as adjusted means rather than observed means. What are they adjusted for? The text of the paper did not mention it nor does the table caption. Please describe this step making sure you justify why adjusted were used and analysed and not observed scores.	The adjusted means are not used in the analysis; they are an output from the ANOVA model. We have added a note to Table 3 to explain that the adjusted means are: "*Estimated marginal means, adjusted to account for between-group differences in the other variables in the model.*"

Tolerance ambiguity in medical students and doctors

Full title; Medical student and junior doctors' tolerance of ambiguity: development of a new scale

Jason Hancock[1], Martin Roberts[2], Lynn Monrouxe[3] & Karen Mattick[1]

1. University of Exeter Medical School, St Luke's Campus, Heavitree Rd, Exeter, EX1 2LU, UK

2. Collaboration for the Advancement of Medical Education Research and Assessment (CAMERA), Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK

3. School of Medicine, Cardiff University, Cardiff, UK

Correspondence address;

Dr. Jason Hancock;

University of Exeter, College House, St Lukes Campus, Heavitree Rd, Exeter, Devon EX1 2LU

j.hancock2@exeter.ac.uk, 01392 725955,

Abstract

The practice of medicine involves inherent ambiguity, arising from limitations of knowledge, diagnostic problems, complexities of treatment and outcome and unpredictability of patient response. Research into doctors' tolerance of ambiguity is hampered by poor conceptual clarity and inadequate measurement scales. We aimed to create and pilot a measurement scale for tolerance of ambiguity in medical students and junior doctors that addresses the limitations of existing scales. After defining tolerance of ambiguity, scale items were generated by literature review and expert consultation. Feedback on the draft scale was sought and incorporated. 411 medical students and 75 Foundation doctors in Exeter, UK were asked to complete the scale. Psychometric analysis enabled further scale refinement and comparison of scale scores across subgroups. The pilot study achieved a 64% response rate. The final 29 item version of the Tolerance of Ambiguity in Medical Students and Doctors (TAMSAD) scale had good internal reliability (Cronbach's alpha 0.80). Tolerance of ambiguity was higher in Foundation Year 2 doctors than first, third and fourth year medical students (-5.23,P = 0.012; -5.98, P = 0.013; -4.62, P = 0.035, for each year group respectively). The TAMSAD scale offers a valid and reliable alternative to existing scales. Further work is required in different settings and in longitudinal studies but this study offers intriguing provisional insights.

Key words;

Ambiguity, epistemology, medical education, tolerance, uncertainty,

Introduction

The practice of medicine involves inherent ambiguity and uncertainty, arising from limitations of knowledge, diagnostic problems, ambiguities of treatment and outcome, and unpredictability of patient response (Geller, Faden, & Levine, D, 1990). The ability of physicians to tolerate ambiguity is therefore of significant interest, with implications for doctors' mental health and wellbeing, staff retention in the medical profession, and specialty choice. For example, low tolerance of ambiguity has been linked with low patient and physician satisfaction, increased risk of physician burnout (Lim, 2003; Cooke, Doust, & Steele, 2013), more negative attitudes towards underserved populations (Wayne, Dellmore, Serna, Jerabek, Timm, & Kalishman, 2011), and personality traits such as dogmatism, conformity and rigidity (Budner, 1962; Furnham, & Ribchester, 1995). It has also been argued that evidence-based complex decision making, which requires the integration of individual patient perspectives and research evidence that may be incomplete, poor quality or conflicting, is only possible if the clinician is able to understand the limits of their own, and of scientific, knowledge and manage the associated uncertainty (Knight, & Mattick, 2006).

It is therefore understandable that undergraduate medical education has been encouraged to introduce educational strategies that will increase medical students' tolerance of ambiguity (Luther, & Crandall, 2011). However, the challenges are formidable. There is a fundamental lack of conceptual clarity around the term 'tolerance of ambiguity' and whether it can change over time in individuals or populations, or what strategies might enable it to change. There may also be unintended consequences associated with increasing all medical students' tolerance of ambiguity (Hancock, & Mattick, 2012). Crucially, the tools available to measure tolerance of ambiguity are crude, despite over 60 years of research, and this provides a particular barrier for the evaluation of educational strategies aimed at increasing learners' tolerance of ambiguity.

This paper aims to seek conceptual clarity around tolerance of ambiguity, to offer a measurement scale that can support the evaluation of educational strategies and to make use of a modern validity assessment framework to evaluate the validity of the scale for use in undergraduate medical students and foundation doctors. It also aims to make some tentative insights into whether the tolerance of ambiguity of populations of students changes during medical school.

Defining ambiguity and uncertainty

The Collins English dictionary defines ambiguity as "vagueness and uncertainty of meaning" and uncertainty as "not known, reliable or definite". These definitions are conceptually similar and are often used interchangeably. However, a distinction has been made by some authors, including Greco & Roger (2002), who suggest that uncertainty is the response to an ambiguous situation, akin to the period of anticipation prior to the confrontation with a potentially harmful event. *Tolerance of ambiguity* has been defined as "the way an individual (or group) perceives and processes information about ambiguous situations or stimuli when confronted by an array of unfamiliar, complex or incongruent clues" (Furnham, & Ribchester, 1995). Indeed, an individual who is intolerant of ambiguity may experience stress when encountering an ambiguous situation, avoid ambiguous stimuli, seek clarity or act prematurely (Furnham, & Ribchester, 1995). *Intolerance of ambiguity* has been defined as the tendency to perceive or interpret ambiguous situations as actual or potential sources of psychological discomfort or threat (Norton, 1975). In these definitions, ambiguity can exist in a situation that is novel, unfamiliar or complex, or when the cues are contradictory. Other authors point out that tolerance of ambiguity might mean not only coping well in ambiguous situations but actively seeking out and thriving in them (Budner, 1962). This suggests the need for a potentially multidimensional construct of tolerance of ambiguity.

For this study, we used the Collins English dictionary definition of ambiguity and consider ambiguity to be the stimulus; and Greco et al.'s definition of uncertainty and consider uncertainty to be the response to an ambiguous situation. Therefore ambiguity and uncertainly are not fully synonymous, with tolerance of ambiguity being more wide ranging than tolerance of uncertainty, although they are closely related. In reality, it is likely that avoidance of uncertainty is correlated with intolerance of ambiguity (Furnham, & Ribchester, 1995). These definitions suggest that tolerance of ambiguity is closely aligned with personal epistemologies, although we are not aware of medical education literature that has explored this interface or overlap. The next section will explore the personal epistemological frameworks used by individuals when learning and making decisions, and the implications for this study.

Personal epistemology and tolerance of ambiguity - state versus trait?

Epistemology is a branch of philosophy that considers what it is to ‘know’: how we understand, integrate, justify, and apply knowledge. Early empirical data and theoretical models suggest that *personal epistemologies* change and develop over time, although there is less consensus about how this happens. Models of personal epistemology therefore describe development from a lay understanding of science, where science is considered to be based on certainties and ‘truths’, to an understanding that is more contextual and fluid (Knight, & Mattick, 2006; Norton, 1975). Research in this field has a long tradition starting with the works of Polanyi (1966) who tells us there is no such thing as objective factual science because all knowledge is understood through our own ‘worldview’, and Perry’s developmental work on the nine stages of maturation (1968). However, in terms of the developmental perspective, early models tend to suggest linear development along an essentially unidimensional scale. More recent models challenge this view, suggesting that personal epistemology can have multiple dimensions, sometimes moving ‘backwards’ along the scale, and is likely to be topic / context specific. For example, in a recent study, first year medical students viewed anatomical knowledge as concrete and certain, whilst accepting that more ambiguity exists within the social sciences. In addition, students also progress to a more contextual and fluid understanding at a different rate for aspects of the same topic (Knight, & Mattick, 2006).

Hammer, and Elby (2002) agree that how an individual makes sense of a situation depends on the context of the situation but suggest individual variation in how the context activates the personal *epistemological resources* at their disposal. Some individuals may make use of a more static framework, based on beliefs surrounding the certainty and unchanging nature of science. Others make use of more dynamic framework, believing that scientific knowledge stems from evidence and changes and expands with everyday life. These frameworks are made up of different epistemological resources that we learn throughout our lives, including beliefs such as knowledge can be accumulated, and knowledge can be checked.

There are a number of implications of the work on personal epistemologies for our study of tolerance of ambiguity. First there is clear overlap of ideas and we would hypothesise that an individual with a more sophisticated personal epistemology would be more tolerant of ambiguity. Secondly, we conclude that we should be open to the possibility that tolerance of ambiguity will change over time, given appropriate environmental conditions and contextual exposure, rather than thinking of tolerance of ambiguity as a stable trait or personality variable as originally proposed by Budner in 1962. In other words, a change in learning context could encourage students to think differently about knowledge and evidence, and / or use different epistemological resources. This opens up exciting possibilities for educational strategies within medical school (Luther, & Crandall, 2011) perhaps involving early clinical and research exposure, and supported reflection. Finally, any attempt to measure a medical student's or doctor's tolerance of ambiguity would need to do so in clinical context, since the cues experienced in this context will activate the individual's epistemological resources and determine the level of ambiguity experienced by the individual practicing doctor.

Implications for measuring tolerance of ambiguity

From this analysis, we propose that a scale to measure tolerance of ambiguity amongst populations of medical students and junior doctors would need to: 1) contain items that are clinically contextualised; 2) have sufficient number and range of items to be sensitive to subtle changes; 3) be open to the possibility that tolerance of ambiguity is a multidimensional construct; and 4) demonstrate good validity evidence (Downing, 2003).

The most widely used scale to date is the original Budner scale (1962) or variations thereof. This scale has 16 items and good construct validity (Sidanius, 1988). However, it conceptualises tolerance of ambiguity as a single dimensional personality measure, the items are not clinically contextualised, and the internal reliability is poor (Cronbach's alpha 0.49 in the original report). Therefore this scale does not meet our four criteria above for a measurement tool for use with medical students. A 4 item modified Budner scale was used by Geller et al. (1990) which contains 3 items taken directly from the original Budner scale. Whilst two of the items were clinically contextualised and the Cronbach's alpha score did increase marginally (0.56), the other reservations still apply. In addition, the small number of items means the scale is unlikely to be sensitive to change. A more recent publication by Geller, Tambor, Chase, & Holtzmann, (1993) introduced a new scale which has an improved internal reliability (Cronbach's alpha 0.75) but is not clinically contextualized, contains only 7 items (reduced from 18 during the pilot study) and is unidimensional. The "Physicians reaction to uncertainty scale" (Gerrity, DeVellis, & Earp, 1990) comes closest to meeting our criteria, being clinically contextualised, containing 61 items and acting as a multidimensional measure. However the focus is very much on practicing physicians, is not validated for use in medical students and seeks to measure reaction to uncertainty rather than tolerance of ambiguity.

Therefore, to our knowledge, no scales exist that meet our requirements and can answer our question about change in tolerance of ambiguity during medical school.

Medical students' tolerance of ambiguity

Little empirical data exist about whether tolerance of ambiguity increases or decreases in medical students during undergraduate education and that which does exist are conflicting and potentially flawed due to the limitations identified in the measurement scales used. Budner (1962) suggests that tolerance of ambiguity may be higher in third year medical students than first years, although this was only found in one of the two medical schools studied and the differences were not statistical significant. In contrast, Geller et al. (1990) showed no difference between levels of ambiguity recorded in medical students (n=86) in their first, second and fourth year of study in a cross-sectional study in one medical school and concluded that selection may be more important than education and training in influencing tolerance of ambiguity amongst physicians. In postgraduate medical training, Deforge, and Sobal (1991) used the original Budner scale with 59 family practice residents in America and found higher tolerance of ambiguity in third year residents compared to first years. Their paper also claimed that medical students were more tolerant of ambiguity than the first year residents, however it is unclear if the 'slightly different' score referred to reached statistical significance. Recently, Geller hypothesised that those students entering medical school with a higher baseline tolerance of ambiguity may show an increase during medical school, while those less tolerant at baseline may show a reduction (Geller, 2013).

The tolerance of ambiguity of different sub-groups has also been explored, albeit with the same poor scales. Geller et al. (1990) reported that female medical students had a higher level of tolerance of ambiguity than male students, while Deforge, and Sobal (1991) found no gender differences in family practice residents. Medical students who were 'older' when they started medical school also had a higher tolerance of ambiguity, although the cut off for this is unclear (Geller et al., 1990).

Similarly, the association between tolerance of ambiguity and choice of medical specialty has been considered and this is perhaps the area of greatest consensus, although again we note the concerns about the data on which these conclusions are made. Budner (1962) reported that first to third year medical students who wished to pursue a 'less structured' career (e.g. psychiatry) were more tolerant of ambiguity; whereas those who wished to pursue the most structured careers (e.g. surgery, obstetrics and gynaecology) were more intolerant of ambiguity. Geller et al. (1990) confirmed that medical students wishing to pursue a career in psychiatry were more tolerant of ambiguity than those wishing to pursue a career in surgery. Furthermore Geller et al. (1993) reported that psychiatrists were more tolerant of ambiguity than obstetricians, paediatricians and family practitioners. Given that the entry criteria for core psychiatry training in the UK include the "capacity to deal with ambiguity & uncertainty in clinical life" then perhaps this is not surprising (Royal College Psychiatrists, 2013).

In summary based on the current literature, which is challenged by the methodological tools available, we can tentatively propose that tolerance of ambiguity may be associated with specialty choice and some demographic markers such as age and gender. It remains unclear whether tolerance of ambiguity changes during medical school but initial evidence suggests it may increase when working as a doctor, albeit a limited range of specialties have been explored.

Study aims and research questions

The aims of this study are twofold;

1. To design a measurement scale for tolerance of ambiguity in medical students and junior doctors that is clinically-contextualised but still relevant for first year medical students, that treats tolerance of ambiguity as a complex construct that may have multiple dimensions and be open to change, and that has a good internal reliability but has sufficient items that is likely to be sensitive to change. To use the results obtained to evaluate the validity of our scale in the population studied.
2. To offer some provisional insights into the associations between tolerance of ambiguity and stage in undergraduate / postgraduate training, demographics such as gender, entry status (e.g. prior degree) and prospective career choices.

The process of evaluating the validity of this scale involves developing a scientifically sound validity argument to support the intended interpretation of test scores for ambiguity and their relevance to the proposed use in the undergraduate medical student and foundation doctor population. The guidelines that we shall follow for this purpose are set out by the American Educational Research Association (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999) and have been applied directly to the medical education context by Downing (2003).

We hypothesise that tolerance for ambiguity will increase during medical education as students gain research skills and clinical experience, coupled with support for reflective practice, that allow them to develop new ways to activate and apply their epistemological resources; and that female participants who are older, have a prior degree on entering medical school, and/or wish to become psychiatrists will have a higher tolerance of ambiguity.

Methods

Item generation

First, a literature review was undertaken to identify the definitions, theories, empirical research and existing scales relevant to tolerance of ambiguity. Eight medical education colleagues at a medical school in South West England were then asked to send us scenarios highlighting examples of ambiguity in medicine and medical education. Using the themes identified from the literature and the scenarios as a guide, 49 draft items were generated (by KM and LM). Each item was placed into one of three subscales depending on what it was measuring; 'tolerance of ambiguity, seeks out and thrives in ambiguous situations', 'tolerance of ambiguity, but seeks to reduce ambiguity', or 'intolerance of ambiguity'. Individual scale items were developed, ensuring that they were clear, short, focused and had good face validity (Oppenheim, 2008). Four of these items were taken directly from the modified Budner scale used by Geller et al. in 1990 and a further 7 were taken from the original Budner scale (1962), although slightly modified to include the medical context. Therefore in total 10 items were taken from the original Bunder scale. All items were written as statements to which respondents were asked to score their agreement on a five-point Likert-type scale ranging from 'strongly disagree' (1) to 'strongly agree' (5), with the midpoint score being 'neutral' (3). This differs from the six point likert scales used previously by both Budner (1962) and Geller et al. (1990). The items focused on tolerance of ambiguity within a medical context, whilst ensuring the content was appropriate for first year medical students, but also relevant for junior doctors.

Item validation

The same eight colleagues reviewed the draft items. They were asked to give their opinion in the form of a numerical rating (1 to 5) on the quality of the item and asked to provide text comments on individual items in terms of (1) their relevance to the construct of tolerance of ambiguity; (2) their clarity; (3) their format; and (4) whether they would be understood by the range of target subjects, from first year undergraduate medical students to Foundation year doctors (junior doctors in their first two years after graduating from medical school). The experts were given the opportunity to reword or remove any items they felt inappropriate. If two experts suggested an item was removed then it was. Through this process, 9 items were removed, 14 more were reworded and 2 new items were written, resulting in an initial set of 42 items.

Ten medical students and Foundation year doctors reviewed the 42-item scale to ensure that the items were clear and understandable. In addition to scoring their agreement with each item, respondents were asked to provide free text comments and to identify poorly worded or difficult to answer items. Following this initial work (results not shown) one item was removed due to poor face validity and 2 more were reworded, resulting in a final 41-item TAMSAD questionnaire (Table 1).

Participants and recruitment

Following ethics approval, 411 undergraduate medical students across years 1-5 and 75 Foundation doctors, all based in Exeter, were approached and asked to complete the TAMSAD questionnaire. Participants were initially approached by e-mail with a link to a short YouTube video filmed by JH. Following this, the research team attended lectures and small group sessions and handed out paper versions of the questionnaire. Any medical student not seen was sent a copy of the TAMSAD by e-mail and asked to participate. All Foundation doctors were provided with electronic versions of the TAMSAD and asked to complete and return electronically. The questionnaire also included demographic questions, entry status (are you a graduate student?) and intentions regarding possible future careers in a range of specialties.

Statistical analysis

Data analysis was conducted in SPSS version 21. We calculated survey response rates within each year group. Negatively worded items were reverse scored prior to analysis. We calculated item response rates and Mahalanobis distances for each respondent and used these to identify, and remove where appropriate, potentially outlying respondents in the data set.

Responses to each item were examined for the distribution of responses, mean response score and standard deviation of response scores. If an item had a mean score below 2 or greater than 4, or if participants did not use the full five-point range of scale responses, then the standard deviation of the scores was examined. If the standard deviation was low, the items were considered for removal from further analysis.

The initial 41-item scale contained the 4 Geller et al. (1990) items and a further 7 items from the original 16 item Budner (1962) scale. We calculated Cronbach's alpha for each of these scales and compared with the values originally reported by those authors. Since we used only 10 of the 16 items from Budner's scale, if we include the 3 also used in the modified Geller scale, we used the Spearman-Brown formula to estimate Cronbach's alpha for the original scale length (Stanley, 1971). Due to the item differences and the incorporation of the items into a longer questionnaire these estimates are only proxies for the reliability of the Budner and Geller et al. scales in our sample.

Exploratory factor analyses using a variety of extraction and rotation procedures were used to investigate the possible existence of subscales within the overall scale. Following this, we developed a final TAMSAD scale by removing items which did not improve the value of Cronbach's alpha for the overall scale. We then removed~~, followed by~~ items for which the adjusted item-total correlation with the remaining items was less than 0.20.

The final TAMSAD scale was used to conduct preliminary analyses on the levels of tolerance of ambiguity of medical students and doctors in the study sample. A TAMSAD score for each respondent was calculated as the mean item score (provided there were no more than two missing items) and linearly transformed from the original 1-5 scale to a 0-100 scale using the formula; New score = 25(Old score -1). The distribution of the scores across all respondents was examined for normality. Analysis of variance (ANOVA) was used to explore the possible influence on the

TAMSAD score of respondents' year group, gender, graduate entry status and interest in any of seven specialities (medicine, surgery, emergency medicine, general practice/ community medicine, psychiatry, paediatrics, radiology). We calculated effect sizes for independent predictors in relation to the magnitude of the standard deviation of the TAMSAD score (Cohen, 1988).

Results

Questionnaire analysis

Three of the 314 returned questionnaires had over half of the items unanswered and were excluded from the analysis. The Mahalanobis distance method identified eight potential outliers. One of these had employed an 'answer 1 or 5 strategy' and this questionnaire was also excluded from the analysis, giving an effective response rate of 310/486 (64%). Response rates varied by year group (Table 2).

Item 31 was removed from the TAMSAD scale because it had the lowest standard deviation of response scores (0.64) and on reflection the researchers felt it could be examining participant 'expertise' rather than their 'tolerance of ambiguity'. Since we wished to investigate the relationship between tolerance of ambiguity and specialty preferences we excluded items 5 and 15, which are clearly linked to such preferences.

The factor analysis indicated that the remaining 38 TAMSAD items could not be subdivided into a simple set of interpretable factors. Using principal factors extraction there were thirteen factors with eigenvalues greater than one but the scree plot suggested a five-factor solution accounting for 33% of the total variance. However the five factors enabled no simple interpretation (even after applying a Varimax rotation) and numerous items either had no factor loadings greater than 0.3 or loaded moderately on more than one factor (Appendix 1). Use of alternative extraction and rotation methods failed to find any simple solution. The initial Cronbach's alpha score was calculated at 0.75 and we interpreted this as suggesting that the TAMSAD questionnaire was acting as a unidimensional measure of tolerance of ambiguity.

We then looked to improve both the parsimony and reliability of the scale by reducing the total number of items. We found that removal of seven items (1, 2, 4, 12, 16, 34, & 40) increased the internal consistency of the scale to 0.80, while reducing the total number of questionnaire items to 31. Finally we removed two further items (14 and 37) which had adjusted item-total correlations less than 0.20. Note that item 11, which had an initial item-total correlation of 0.19 (Table 1), was retained at this stage as its item-total correlation with the remaining items was now above 0.20. This left the Cronbach's alpha unchanged at 0.80 indicating that the scale has a good internal consistency (Field, 2005) and could be interpreted as a unidimensional measure.

Three of these items had originally been intended to measure 'tolerance of ambiguity but seeks to reduce ambiguity' (4, 12, 16) rather than purely 'tolerance of ambiguity, seeks out and thrives in ambiguous situations'. Therefore it is not surprising that their removal from a scale that appears to be acting in a unidimensional way improves the internal consistency of the scale.

Table 1 shows items means and standard deviations for the original 41 items, indicates those which items came from the original Budner and Geller et al. scales, those which were retained in the final TAMSAD scale and their corrected item-total correlations with the scale.

We estimated the internal consistency reliability of the Geller et al. 4 item scale to be 0.31 and that of the full Budner 16 item scale to be 0.63.

Group differences

Participant scores on the 29-item TAMSAD scale ranged from 38.8 to 86.2 with a mean (SD) of 57.0 (8.8). Using the TAMSAD scale we found that significant differences in tolerance of ambiguity were associated with participants' year group but not with gender, graduate entry status or possible future career specialty interests (Table 3). Observed mean TAMSAD scores by year group are shown in Figure 1. First, third and fourth year medical students had significantly lower tolerance of ambiguity than FY2 doctors by 4.62 to 5.98 scale points. These are moderate effects (0.52 to 0.68 times the TAMSAD score SD) (Cohen, 1988). Tolerance of ambiguity in second and fifth year medical students and FY1 doctors was similar to that in FY2 doctors.

Participants expressing a preference for a possible career in surgery were, on average, 2.52 scale points lower in their tolerance of ambiguity than their peers, while those preferring paediatrics were 2.42 points higher but neither of these differences reached statistical significance at the 0.05 level. Prospective medics, GPs, emergency physicians, psychiatrists and radiologists had equivalent tolerance of ambiguity to their peers.

Discussion

This study aimed to design a scale to measure tolerance of ambiguity in medical students and junior doctors that address the limitations of existing scales, and to pilot it with the target population in one location. After several rounds of refinement, we arrived at a 29-item scale that we named the Tolerance of Ambiguity in Medical Students and Doctors (TAMSAD) scale.

We evaluated the validity of this scale using an established framework set out by the American Educational Research Association (AERA, APA, & NCME, 1999) and applied to the medical education context by Downing (2003). This framework states that when evaluating the validity of any assessment tool used in medical education five sources of evidence should be considered; content related validity evidence, the response process, the internal structure of the scale, the relationship to other variables and the consequences of using the assessment scale.

Content related validity evidence was provided through the provenance of the items, which were derived from an analysis of the education literature, from medical education theory and from existing tolerance of ambiguity scales. Since we did not want to assume that our scale would be acting as a unidimensional measure of ambiguity, items were initially separated into one of three subscales: ‘tolerance of ambiguity, seeks out and thrives in ambiguous situations’, ‘tolerance of ambiguity, but seeks to reduce ambiguity’, and ‘intolerance of ambiguity’. Pilot work involved the input of academic staff working in medical education and from medical practitioners working in hospital and community settings. Unlike many previous scales, TAMSAD is context specific which allows it to assess an individual’s tolerance of ambiguity in the medical setting. The pilot study achieved a 64% response rate across the five years of medical students and two years of Foundation doctors.

The response process was considered in the scale development stage as academics and clinicians were asked to remove or reword items that they felt inappropriate or difficult to understand. A pilot study was also completed during which 10 medical students and foundation doctors were asked to comment if they felt that items within the scale were difficult to understand or answer. Finally we ensured that data collected from the scales was accurately transcribed onto the statistical package by performing a thorough check of the data.

The internal structure of the TAMSAD scale was explored by measuring its internal reliability (Cronbach's alpha=0.80). We have interpreted this to mean that the scale is acting as a unidimensional measure of tolerance of ambiguity, which is supported by the improvement in internal consistency observed when we removed three of the four items initially created to measure 'tolerance of ambiguity, but seeks to reduce ambiguity'. The Cronbach's alpha associated with the TAMSAD scale in this population was higher than obtained by previous scales; 0.49 (Budner, 1962) and 0.56 (Geller et al., 1990). The Cronbach's alpha associated with the original Geller et al. questionnaire was lower than reported previously in our study population (0.31 in our study compared to 0.56). The expected Cronbach's alpha for the full Budner 16 item scale was 0.63, which was higher than Budner's reported value of 0.49, perhaps due to sample heterogeneity (i.e. if the participants in our sample were more varied in their tolerance of ambiguity than were those in Budner's sample then we would expect reliability to be higher).

The relationship of tolerance of ambiguity (as measured by TAMSAD) and other variables, such as stage of training, gender, graduate status and specialty choice, was sought. In our study, foundation year 2 doctors had a higher tolerance of ambiguity than first, third and fourth year medical students; other studies have reached various conclusions about association of tolerance of ambiguity and stage of training. Our study was a cross sectional survey, so conclusions about changes in tolerance of ambiguity over time cannot be made. Wayne et al. (2011) demonstrated that those students with a higher tolerance of ambiguity at the start of medical school showed a smaller deterioration in their attitudes towards underserved populations and other studies have shown that intolerance of ambiguity is associated with distress (Benbassat, Baumal, Chan, &Nirel, 2011) and reduced levels of work satisfaction (Bovier, & Perneger, 2007). In response, Luther et al. (2011)

recommended that medical schools do more to increase their students' tolerance of ambiguity but a responding commentary cautioned about the possible unintended consequences of doing this, for example resulting in an undersupply of surgical trainees. In reality it is likely that there are multiple factors underpinning any increase in tolerance of ambiguity during medical school, including the increasing maturity of students. Our study showed no significant association between tolerance of ambiguity in prospective surgeons or psychiatrists compared with their peers. Previous research has suggested an association between tolerance of ambiguity and medical career intention (Budner, 1962; Geller, Tambor, Chase, & Holtzmann, 1993).

We would argue that the consequences of completing this questionnaire are minimal. Our research suggests that completing the modified twenty-nine questionnaire will take 5–10 minutes and is unlikely to have a negative impact on participants.

The most important contribution of this study is to provide a valid tool for the research community to apply in subsequent studies. The provisional findings from piloting the scale are broadly supportive of previous research by Budner (1962) and Geller et al. (1990).

As with all studies, our work has a number of methodological strengths and challenges. The strengths of this study include the way in which we carefully defined the constructs of ambiguity and uncertainty, thus ensuring a rigorous process in the development of the scale. Additionally, there were multiple rounds of scale refinement based on target group feedback and on psychometric analysis. The pilot study achieved a large sample size and good response rate. In terms of challenges, data collection only took place in one site, and the study was cross sectional in design rather than longitudinal. These both serve to limit the conclusions that can be drawn in relation to change over time.

Future research

Further work to provide additional evidence for the validity of the TAMSAD scale is now required. One aspect that has not been explored in depth is the cultural sensitivity of the scale. The current study has used the scale in one location in South West England; future work could use the scale in different countries and with more culturally heterogonous populations. The fact that the scale seemed to be acting as a unidimensional measure was unexpected, given the theoretical complexity of the construct of tolerance of ambiguity. It would therefore be helpful to repeat the exploratory factor analysis process with other, larger populations of students to verify this observation. Further qualitative research could also be useful to explore the different aspects of the construct of tolerance of ambiguity in different settings and with different populations.

It will also be interesting to explore associations between tolerance of ambiguity (as measured by TAMSAD) with other variables such as attitudinal markers (e.g. cynicism), observed behaviours (e.g. medical professionalism), cognitive states (e.g. intellectual development) and aspiration (e.g. specialty choice). Specific variables of interest could include attitudes towards underserved populations (Wayne, 2011) and intellectual maturation (Perry, 1968).

Other work should develop the TAMSAD for use in an online forum. We were able to ensure good response process validity as the scale was only completed on a small scale, mainly on paper. If the scale is to be used on a larger population, then the use of an electronic scale would help to ensure that there are minimal errors when processing data.

Finally, following further validation, this scale could be used in a number of different ways to shed light onto how tolerance for ambiguity might change during medical undergraduate and early postgraduate training. Whilst our cross sectional survey design has provided interesting data, future research involving longitudinal methodologies would enable us to track students through medical school and into early clinical practice, therefore providing insights into the pattern of growth or decline in tolerance of ambiguity during medical school and junior doctor training. Such studies might enable us to draw conclusions about how levels of tolerance of ambiguity vary across different medical schools using different curricula.

Conclusion

The TAMSAD scale developed through this study offers a more valid and reliable alternative to existing scales for medical students and junior doctors. Further work is now required to continue the process of evaluating the validity of this scale in the undergraduate and foundation doctor population. This will be possible through conducting longitudinal studies to explore changes in tolerance of ambiguity, both over time and as a result of educational interventions. Meanwhile this study offers intriguing provisional insights that warrant further investigation.

Acknowledgements

We would like to thank everybody who contributed towards each stage of this study. Our eight medical education colleagues who contributed towards initial item creation, the ten medical students and foundation doctors in Exeter who took part in the initial study pilot and the medical students and foundation doctors in Exeter who completed the TAMSAD questionnaire.

References;

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association, 1999.

Benbassat, J. Baumal, R. Chan, S. &Nirel, N. (2011). Sources of distress during medical training and clinical practice: Suggestions for reducing their impact. *Medical Teacher*, 33(6), 486-490.

Bovier, P. &Perneger, T. (2007). Stress from uncertainty from graduation to retirement – a population-based study of Swiss physicians. *Journal of General Internal Medicine*, 22(5), 632-638.

Budner, S. (1962). Intolerance of ambiguity as a personality variable. *Journal of Personality*, 30(1), 29-50.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences.* (Lawrence Erlbaum Associates).

Cooke, G. Doust, J. & Steele, M. (2013). A survey of resilience, burnout, and tolerance of uncertainty in Australian general practice registrars. *BMC Medical Education*, 13, 2.

Downing, S. (2003). Validity: on the meaningful interpretation of assessment data. *Medical Education*, 37, 830-837.

Field A. (2005). *Discovering statistics using SPSS*, Second edition. (Sage publications).

Furnham, A. & Ribchester, T. (1995). Tolerance of ambiguity: A review of the concept, its measurement and applications. *Current Psychology*, 14(3), 179.

Geller, G. Faden, R. & Levine, D. (1990). Tolerance for ambiguity among medical students: implications for their selection, training and practice. *Social Science & Medicine*, 31(5), 619-624.

Geller, G. Tambor, E. Chase, G. &Holtzmann, N. (1993). Measuring physicians' tolerance for ambiguity and its relationship to their reported practices regarding genetic testing. *Medical Care*, 31(11), 989-1001.

Geller, G. (2013). Tolerance for ambiguity: An ethics-based criterion for Medical Student selection, *Academic Medicine*, 88(5), 581-584.

Gerrity, MS. DeVellis, RF. Earp, JA. (1990). Physicians' reactions to uncertainty in patient care. A new measure and new insights. *Medical Care*, 28(8), 724-736.

Greco, V. & Roger, D. (2002) Uncertainty, stress and health. *Personality and Individual Differences*, 34(6), 1057-1068.

Hammer, D. & Elby, A. (2002). On the form of a Personal Epistemology. (In Hofer, B. &Pintrich, P. (Eds.), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing* (pp169-190). Mahwah, NJ: Erlbaum.).

Hancock, J. & Mattick, K. (2012). Increasing Students' Tolerance of Ambiguity: The Need for Caution, Letters to the Editor. *Academic Medicine*, 87(7), 834.

Knight, L. & Mattick, K. (2006). When I first came here, I thought medicine was black and white: making sense of medical students' ways of knowing. *Social Science & Medicine*, 63(4), 1084-96.

Lim, M. (2003, April). Who Is Being Difficult? Addressing the Determinants of Difficult Patient-Physician Relationships.Virtual Mentor, 5, 4. Retrieved 13 June 2013, from http://virtualmentor.ama-assn.org/2003/04/jdsc2-0304.html

Luther, VP. & Crandall, SJ. (2011). Commentary: ambiguity and uncertainty: neglected elements of medical education curricula? *Academic Medicine*, 86(7), 799-800.

Norton, R. (1975). Measurement of ambiguity tolerance. *Journal of Personality Assessment*, 39(6), 607–619.

Oppenheim, AN. (2008). *Questionnaire design, interviewing and attitude measurement*. (London: Continuum).

Perry, W. (1968). *Forms of Intellectual and Ethical Development in the College Years – A Scheme*. (New York: Holt, Rinehart and Winston).

Polanyi, M. (1966). *The Tacit Dimension*. (London: Routledge).

Royal College Psychiatrists CT1 personal specification. Retrieved 24 March 2013 from;

http://www.rcpsych.ac.uk/pdf/CT1%20Psych-person%20spec-August%202013-approved.pdf

Sidanius, J. (1988). Intolerance of ambiguity, conservatism and racism: Whose fantasy, whose reality? A reply to Ray. *Political Psychology*, 9, 309-316.

Sobal, J. &,Deforce, B. (1992). Reliability of Budner's intolerance of ambiguity scale in medical students. *Psychological Reports*, 71, 15-18.

Stanley, J. (1971). *Educational Measurement*. Second edition. (Washington, DC: American Council on Education).

TAMSAD recruitment video. Retrieved 13 June 2013 from;

http://www.youtube.com/watch?v=uQEANkFzfq4&feature=youtu.be

Wayne, S. Dellmore, D. Serna, L. Jerabek, R. Timm, C. & Kalishman, S. (2011). The association between intolerance of ambiguity and decline in medical students attitudes toward the underserved. *Academic Medicine*, 86(7), 877-882.

Figure 1. Observed mean tolerance of ambiguity score (with 95% confidence interval) by year group. FY = Foundation year.

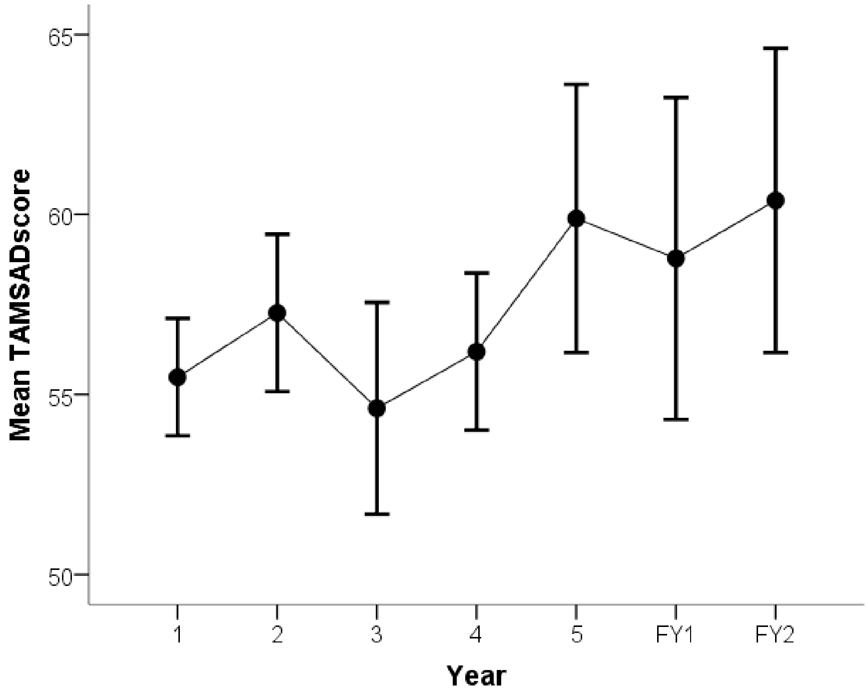


Table1; The TAMSAD scale.

	Item	Mean (likert score)	SD (likert score)	Was the item included in the final 30 item scale?	Corrected item-total correlation ~~(all 41 items)~~	~~Corrected item-total correlation (final scale)~~
1	I am comfortable to acknowledge that I'll never know everything about medicine.	4.28	.97	No	.13	~~n/a~~
2	Even when there is conflicting information, I prefer to make a decision and move on.*	3.27	.94	No	.01	~~n/a~~
3	I would enjoy tailoring treatments to individual patient problems.	4.11	.69	Yes	.30	~~.26~~
4	I think it is important to attribute a percentage likelihood to a diagnosis or a specific patient outcome.	3.45	.91	No	-.05	~~n/a~~
5	As a doctor I would prefer the clear and definite work of someone like a surgeon to the uncertainties of a psychiatrist.*,$^{\alpha}$	3.14	1.24	No	.30	~~n/a~~
6	I have a lot of respect for consultants who always come up with a definite answer.*	3.65	.91	Yes	.40	~~.35~~
7	I would be comfortable if a clinical teacher set me a vague assignment or task.$^{\pi}$	2.46	1.01	Yes	.30	~~.31~~
8	A good clinical teacher is one who challenges your way of looking at clinical problems.$^{\pi}$	4.31	.67	Yes	.27	~~.30~~
9	What we are used to is always preferable to what is unfamiliar.*,$^{\alpha,\pi}$	3.16	1.13	Yes	.24	~~.24~~

10	I feel uncomfortable when people claim that something is ‘absolutely certain’ in medicine.	3.51	1.01	Yes	.24	~~.24~~
11	A doctor who leads an even, regular work life with few surprises, really has a lot to be grateful for.*,π	2.91	1.01	Yes	.19	~~.23~~
12	I enjoy reducing the complexity of medical information to something more tangible.	3.95	.80	No	-.08	~~n/a~~
13	I think in medicine it is important to know exactly what you are talking about at all times.*	3.22	1.03	Yes	.32	~~.31~~
14	‘I don’t know’ are really important words in medicine.	4.19	.81	No	.18	~~n/a~~
15	I would prefer to work in a medical specialty where patients normally get better after treatment.*	3.92	.85	No	.24	~~n/a~~
16	I enjoy reducing detailed scientific problems to their core concepts.	3.91	.84	No	-.02	~~n/a~~
17	I feel comfortable that in medicine there is often no right or wrong answer.	3.64	.97	Yes	.41	~~.38~~
18	A patient with multiple diseases would make a doctor’s job more interesting.	3.51	.87	Yes	.30	~~.29~~
19	I am uncomfortable that a lack of medical knowledge about some diseases means we can’t help some patients.*	3.49	.95	Yes	.24	~~.23~~
20	The unpredictability of a patient’s response to medication would bring welcome complexity to a doctor’s role.	2.84	.88	Yes	.24	~~.25~~
21	It is important to appear knowledgeable to	3.37	.99	Yes	.34	~~.32~~

	patients at all times.*					
22	Being confronted with contradictory evidence in clinical practice makes me feel uncomfortable.*	3.03	.93	Yes	.41	~~.43~~
23	I like the mystery that there are some things in medicine we'll never know.	3.20	1.13	Yes	.34	~~.34~~
24	Variation between individual patients is a frustrating aspect of medicine.*	2.20	.95	Yes	.43	~~.42~~
25	I find it frustrating when I can't find the answer to a clinical question.*	3.82	.80	Yes	.40	~~.37~~
26	I am apprehensive when faced with a new clinical situation or problem.*,$^{\pi}$	3.10	1.00	Yes	.31	~~.37~~
27	I feel uncomfortable knowing that many of our most important clinical decisions are based upon insufficient information.*	2.95	.86	Yes	.32	~~.33~~
28	No matter how complicated the situation, a good doctor will be able to arrive at a yes or no answer.*,$^{\pi}$	2.30	.93	Yes	.28	~~.26~~
29	I feel uncomfortable when textbooks or experts are factually incorrect.*	3.76	.92	Yes	.24	~~.21~~
30	There is really no such thing as a clinical problem that can't be solved.*,$^{\pi}$	2.22	.86	Yes	.30	~~.29~~
31	It's an exciting feeling when you listen to a patient tell you their symptoms and you just know what disease it is.*	1.6	.64	No	.06	~~n/a~~
32	I like the challenge of being thrown in the deep end with different medical situations.	3.52	.88	Yes	.33	~~.42~~
33	It is more interesting to tackle a complicated	3.58	.96	Yes	.24	~~.26~~

	clinical problem that to solve a simple one.[π]					
34	In medicine as in other professions, it is possible to get more done by tackling small, simple problems rather than large and complicated ones.*[α,π]	3.33	.85	No	.00	~~n/a~~
35	I enjoy the process of working with a complex clinical problem and making it more manageable.	3.88	.67	Yes	.20	~~.21~~
36	A good job is one where what is to be done and how it is to be done are always clear.*[α,π]	2.98	1.04	Yes	.40	~~.39~~
37	Medicine has a lot of grey areas because we haven't found the answers yet.*	4.11	.83	No	.18	~~n/a~~
38	To me, medicine is black and white.*	1.58	.73	Yes	.28	~~.26~~
39	The beauty of medicine is that it's always evolving and changing.	4.27	.70	Yes	.35	~~.39~~
40	I enjoy working out which opinion is right in situations where many different opinions are expressed.*	3.69	.77	No	-.19	~~n/a~~
41	I would be comfortable to acknowledge the limits of my medical knowledge to patients.	4.01	.86	Yes	.27	~~.27~~

* *Negatively worded items that were reverse scored prior to analysis.*

[π] *Taken from the original Budner (1962) scale.*

[α] *Taken from the Geller et al. (1990) scale.*

Table 2, Response rates by year group.

Stage / year of training	Response rate	%
Undergraduate		
1	74/ 110	67%
2	72/112	64%
3	34/72	47%
4	52/78	67%
5	30/39	77%
Foundation training		
F1	22/39	56%
F2	26/36	72%
All	310/486	64%

Table 3. ANOVA results. Dependent variable: tolerance of ambiguity score.

Factor	Difference in TAMSAD score	P-value	LCL [a]	UCL [b]	Adjusted mean TAMSAD score [c]
Gender					
Female	1.03	0.329	-1.04	3.09	59.63
Male	Reference category		-	-	58.60
Entry status					
Graduate entry	2.27	0.168	-0.96	5.49	60.25
Non-graduate entry	Reference category		-	-	57.98
Year group		0.044 [~~c~~d]			
Year 1	-5.23	0.012	-9.33	-1.14	57.11
Year 2	-3.87	0.065	-7.98	0.24	58.47
Year 3	-5.98	0.013	-10.67	-1.28	56.36
Year 4	-4.62	0.035	-8.92	-0.33	57.72
Year 5	-0.08	0.973	-4.77	4.61	62.26
FY1 [~~d~~e]	-2.80	0.277	-7.87	2.26	59.54
FY2 [~~d~~e]	Reference category		-	-	62.34
Specialty preferences [~~e~~f]					
Medicine	0.92	0.387	-1.17	3.02	59.57
Surgery	-2.52	0.055	-5.09	0.05	57.85
Emergency	1.46	0.214	-0.85	3.77	59.84
GP / community	-1.18	0.287	-3.35	0.99	58.52
Psychiatry	2.22	0.233	-1.44	5.89	60.23
Paediatrics	2.42	0.062	-0.12	4.96	60.32
Radiology	-0.96	0.793	-8.19	6.27	58.63

a. LCL = Lower 95% confidence limit for the difference in score

b. UCL = Upper 95% confidence limit for the difference in score

~~b.~~c. Estimated marginal means, adjusted to account for between-group differences in the other variables in the model

~~c.~~d. P-value for F test of year group as a factor

~~d.~~e. FY = Foundation year

~~e.~~f. Specialty preferences were not exclusive. The reference category is ‘No interest in the specialty’ in all cases. Adjusted means for respondents not interested in each specialty can be obtained by subtracting the difference given in column 2 from the adjusted mean given in column 6.

Appendix 1. Factor analysis.

Rotated Factor Matrix (Loadings <0.1 suppressed)					
	Factor				
	1	2	3	4	5
Q38	.555	.107			
Q28	.518			.115	.170
Q21	.435		.204		
Q06	.433		.155	.101	.242
Q30	.429				
Q14	.391			-.250	
Q36	.373	.205	.140		.148
Q13	.371		.237		.142
Q41	.338		.160		-.150
Q03	.298	.261			-.151
Q01	.253		.231		-.230
Q10	.206	.106	.105	.102	
Q33		.629			
Q08	.126	.510			-.160
Q18		.506	.116		.121
Q35		.437			-.176
Q39	.144	.423	.153	.213	
Q24	.295	.369	.115	.145	.146
Q25			.548	.137	.181

Q23	.111		.499		
Q17	.256	.100	.424	.122	-.123
Q19			.356		
Q20		.200	.315		.180
Q29			.292		.184
Q32		.379	.166	.522	
Q11				.465	.232
Q26		.237	.163	.445	
Q22	.204		.284	.436	
Q40		-.313		-.327	.122
Q07			.249	.315	
Q27	.159		.289	.315	
Q02				-.215	.477
Q12					-.403
Q37			.101		.294
Q16		.199	-.109		-.289
Q34					.268
Q09	.117	.163		.157	.236
Q04					-.153

Extraction Method: Principal Axis Factoring.

Rotation Method: Varimax with Kaiser Normalization.