Assessing teamwork attitudes in healthcare: development of the TeamSTEPPS teamwork attitudes questionnaire

David P Baker,1 Andrea M Amodeo,2 Kelley J Krokos,2 Anthony Slonim,3 Heidi Herrera4

ABSTRACT

Introduction The report, To Err is Human, indicated that a large number of deaths are caused by medical error. A central tenet of this report was that patient safety was not only a function of sophisticated healthcare technology and treatments, but also the degree to which healthcare professionals could perform effectively as teams. Research suggests that teamwork comprises four core skills: Leadership, Situation Monitoring, Mutual Support and Communication. In healthcare, team training programmes, such as TeamSTEPPS®, are designed to improve participant knowledge of, attitudes towards, and skills in these core areas. If such training programmes are effective, changes in knowledge, attitudes and skills should be observed. The purpose of this study was to develop and validate the TeamSTEPPS Teamwork Attitudes Questionnaire (T-TAQ), a measure designed to assess attitudes towards the core components of teamwork in healthcare.

Method A pilot test version of the survey was developed and administered to 495 respondents from various healthcare organisations.

Results Classical item statistics were used to select the final T-TAQ items. Based on this analysis, 30 of the original 110 items were selected for inclusion in the final instrument. Scale reliabilities exceed 0.7, and scales were found to be moderately correlated.

Discussion The T-TAQ provides a useful, reliable and valid tool for assessing individual attitudes related to the role of teamwork in the delivery of healthcare. Issues related to its use and interpretation are discussed.

INTRODUCTION

Public reaction to the problem of patient safety reached a critical mass with the publication of To Err is Human, which concluded that medical errors cause up to 98 000 deaths annually.1 This report generated a demand for action that was headed by the government, the media and the healthcare community. A central tenet was that safety was not only a function of sophisticated healthcare technology and treatments but also the degree to which healthcare professionals perform effectively as teams. The delivery of care by its very nature requires doctors, nurses and other allied health professionals to coordinate their actions.

There is general agreement in the literature that a ‘team’ consists of two or more individuals with specialised knowledge; have specific roles, make decisions, perform interdependent tasks and are adaptable; and share a common goal such as safe care for patients.2 Team performance has been described in terms of classical systems theory, which posits that team inputs, team processes and team outputs are arrayed over time. Inputs include the characteristics of the task to be performed, the elements of the context in which work occurs and the attitudes team members bring to a team situation. Process consists of the interaction and coordination requirements that must occur among team members if a team is to achieve its specific goals. Outputs consist of the products that result from team process. Teamwork occurs in the process phase, during which team members interact and coordinate.3

Numerous models exist, describing the elements of team process. Salas and colleagues reviewed this research and proposed that teamwork consists of the following core competencies: leadership, mutual performance monitoring, backup behaviour, adaptability and team orientation. These core competencies are supported by mutual trust, closed-loop communication, and shared mental models. Salas et al argued that these competencies should be the basis for team training programmes.4

Relying on the Salas review5, Alonso, Baker and colleagues hypothesised that four teamwork skills are critical in the safe delivery of care: leadership, mutual support, situation monitoring and communication. When learnt by team members, these skills produce performance-based, knowledge-based and attitudinal outcomes. For example, shared mental models are viewed to result from the team leader’s specification of a plan of care and identification of individual team member roles and responsibilities. Table 1 presents the core teamwork skills. Healthcare team training programmes, most notably TeamSTEPPS (Team Strategies and Tools to Enhance Performance and Patient Safety), have been designed to target these skills.5

TeamSTEPPS was released by the Agency for Healthcare Research and Quality (AHRQ) in 2006 as a public domain resource kit for teaching specific tools and strategies to healthcare professionals in order to enhance core teamwork skills. TeamSTEPPS also includes in-depth instruction on how to change organisational culture based upon the work of Kotter. Using an action planning change model, TeamSTEPPS involves an assessment phase, a planning phase, an implementation phase, and a sustainment phase.6 The objective is to customise the TeamSTEPPS tools to specific needs to improve team performance and care quality.

TeamSTEPPS has gained a significant amount of traction in a very short time; however, little...
research exists on the effectiveness of team training on improving patient safety. This is important so that research can uncover the relation between team training, improved performance and better patient outcomes. A comprehensive meta-analysis by Salas et al reported that team training programmes can improve team process by upwards for 20%; however, out of the 10 000 teams represented in the database only 181 were healthcare teams. Therefore, questions remain about the true efficacy of team training in healthcare. Researchers are limited in conducting such investigations, because few tools are available to measure changes in teamwork knowledge, attitudes, and skills.

Kirkpatrick’s model of training evaluation is the most widely used approach to assess training effectiveness.3 Kirkpatrick’s model comprises four levels: (1) trainee reactions, (2) trainee learning, (3) transfer and (4) organisational outcomes. This model has remained robust over the years with few changes being advocated. Notably, Kraiger and colleagues decomposed trainee learning into cognitive, affective and skill-based outcomes. Kraiger et al argued that such a conceptualisation provides a more precise understanding of how learning occurs.

Given the need to understand the effects of team training programmes such as TeamSTEPPS on team performance in healthcare, the purpose of this effort was to develop a reliable and valid measure of healthcare professionals’ attitudes towards teamwork. A search of the literature revealed that few such measures exist, particularly oriented towards healthcare, and none were aligned with what the literature advocates as the core components of teamwork. For example, in aviation, the Cockpit Management Attitudes Questionnaire assesses leadership, coordination and communication. In healthcare, the Safety Climate Survey measures perceptions of organisational commitment to patient safety through constructs such as commitment to safety, leadership, interpersonal interactions, attitudes towards stress and knowledge of how to report adverse events, and the Safety Attitudes Questionnaire measures attitudes about teamwork climate, safety climate, perceptions of management, job satisfaction, working conditions and stress recognition.

Although appropriate for the purposes for which they were designed, these measures do not capture attitudes towards what research suggests are the core concepts of teamwork. To address this need, the purpose of this effort was to develop and establish the psychometric properties of the TeamSTEPPS Teamwork Attitudes Questionnaire (T-TAQ). Although the T-TAQ was developed to be aligned with TeamSTEPPS, the T-TAQ can be used to assess healthcare professionals’ attitudes towards the core components of teamwork defined by Alonso, Baker and colleagues. This research is important, because sound tools are required to evaluate the effectiveness of TeamSTEPPS and similar programmes. Such evaluation cannot rely solely on trainee reactions; it must be comprehensive and assess multiple levels of Kirkpatrick’s hierarchy.

METHODS
Item development
Items were developed through an extensive item-writing process that included multiple item writers experienced in survey development and knowledgeable about the principles of teamwork. Items were linked to specific TeamSTEPPS constructs, including leadership, situation monitoring, mutual support and communication. In addition, the construct of team structure was included. In TeamSTEPPS, team structure is a core module within the curriculum. In the broader research domain, team structure is a critical input variable that can influence team process. Item-writing resulted in 110 items.

Next, items were reviewed to edit and narrow the existing item pool. To prevent social desirability, phrases such as ‘it is important to’ were excluded from items, because this may trigger a positive response (ie, agreement with the item), regardless of the true attitude of the respondent. Items were also reviewed to ensure that they asked about the respondent’s attitude and not that of other team members or the team as a whole. Following the review, 64 items remained in the pool (refer to table 2).

Participants
The pilot version of the T-TAQ was distributed to military healthcare providers between December 2007 and April 2008 and participants at a Mid-Atlantic critical care conference in Spring 2008 (ie, civilian sample). Within the military sample, the T-TAQ was completed either prior to or immediately following TeamSTEPPS training. Within the civilian sample, the T-TAQ was administered during the conference.

Four hundred and ninety-five respondents completed the T-TAQ, n=346 military and n=149 civilian. The data obtained from the two samples included different demographic items, so that civilian and military trainers could collect customised data to meet their particular needs. Therefore, common demographic information, which would allow comparisons between the samples, was limited.

RESULTS
Data quality
Analyses were conducted to remove cases with excessive missing data and anomalous response patterns. Cases in which respondents did not respond to at least 95% of the survey items or answered at least 95% with the same response were removed. When respondents provided multiple responses to an item, or the response was unclear, the data point was coded as ‘missing.’

<table>
<thead>
<tr>
<th>Table 2 Number of pilot test items by Team Strategies and Tools to Enhance Performance and Patient Safety construct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construct</td>
</tr>
<tr>
<td>Team structure</td>
</tr>
<tr>
<td>Leadership</td>
</tr>
<tr>
<td>Situation monitoring</td>
</tr>
<tr>
<td>Mutual support</td>
</tr>
<tr>
<td>Communication</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
This resulted in a loss of 46 (9.8%) cases; reducing the sample size from 495 to 449.

Pilot test sample
Of the final 449 cases included for analysis, 311 were from the military, and 138 were from the civilian participants. The majority of these respondents (n=408) reported that they deliver direct care to patients. Of these, 175 reported that they deliver inpatient care, 89 reported that they deliver outpatient care, and 151 reported that they deliver both inpatient and outpatient care. A total of 211 respondents reported their position as physicians (ie, MD or DO) or dentists. Eighty-five were registered nurses (RNs) and 27 advanced practice nurses (APNs). The majority of respondents (n=260) reported working predominantly in intensive care. Eighteen respondents reported working in an Emergency Department and an equal number in a Paediatric Unit. The majority (n=242) reported they had not had prior team training.

Item selection and final questionnaire
Classical item statistics were used to select the final T-TAQ items. First, means, item–total correlations, and standard deviations were computed. Items were selected for inclusion in the final scale if (a) their item–total correlation was 0.30 or higher and (b) including the item in the final T-TAQ improved the scale’s reliability. Finally, to achieve five subscales of equal length, the number of items per subscale was considered. This process resulted in one item being included in the Team Structure subscale whose item-subscale correlation was below 0.50 (r=0.26). In total, 34 items were deleted.

The final questionnaire (presented in online Appendix A) includes 30 items; six items measuring each construct. Among the final items, four are reverse-coded: three in Mutual Support (items 20, 21 and 24) and one in Communication (item 30). Final constructs and their associated scale reliabilities are provided in table 3. We also examined construct independence by intercorrelating the four T-TAQ subscales. Table 4 provides these results. Coefficients ranged from 0.36 (Mutual Support and Team Structure) to 0.63 (Situation Monitoring and Communication). These results suggest that while the constructs overlap to some degree, they also assess unique variance.

Table 3  Team Strategies and Tools to Enhance Performance and Patient Safety Teamwork Attitudes Questionnaire reliability coefficients

<table>
<thead>
<tr>
<th>Construct</th>
<th>No of survey items</th>
<th>Cronbach alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team structure</td>
<td>6</td>
<td>0.70</td>
</tr>
<tr>
<td>Leadership</td>
<td>6</td>
<td>0.81</td>
</tr>
<tr>
<td>Situation monitoring</td>
<td>6</td>
<td>0.83</td>
</tr>
<tr>
<td>Mutual support</td>
<td>6</td>
<td>0.70</td>
</tr>
<tr>
<td>Communication</td>
<td>6</td>
<td>0.74</td>
</tr>
</tbody>
</table>

This resulted in a loss of 46 (9.8%) cases; reducing the sample size from 495 to 449.

Pilot test sample
Of the final 449 cases included for analysis, 311 were from the military, and 138 were from the civilian participants. The majority of these respondents (n=408) reported that they deliver direct care to patients. Of these, 175 reported that they deliver inpatient care, 89 reported that they deliver outpatient care, and 151 reported that they deliver both inpatient and outpatient care. A total of 211 respondents reported their position as physicians (ie, MD or DO) or dentists. Eighty-five were registered nurses (RNs) and 27 advanced practice nurses (APNs). The majority of respondents (n=260) reported working predominantly in intensive care. Eighteen respondents reported working in an Emergency Department and an equal number in a Paediatric Unit. The majority (n=242) reported they had not had prior team training.

Item selection and final questionnaire
Classical item statistics were used to select the final T-TAQ items. First, means, item–total correlations, and standard deviations were computed. Items were selected for inclusion in the final scale if (a) their item–total correlation was 0.30 or higher and (b) including the item in the final T-TAQ improved the scale’s reliability. Finally, to achieve five subscales of equal length, the number of items per subscale was considered. This process resulted in one item being included in the Team Structure subscale whose item-subscale correlation was below 0.50 (r=0.26). In total, 34 items were deleted.

The final questionnaire (presented in online Appendix A) includes 30 items; six items measuring each construct. Among the final items, four are reverse-coded: three in Mutual Support (items 20, 21 and 24) and one in Communication (item 30). Final constructs and their associated scale reliabilities are provided in table 3. We also examined construct independence by intercorrelating the four T-TAQ subscales. Table 4 provides these results. Coefficients ranged from 0.36 (Mutual Support and Team Structure) to 0.63 (Situation Monitoring and Communication). These results suggest that while the constructs overlap to some degree, they also assess unique variance.

Finally, we tested for differences in attitudes across subgroups by conducting t tests using the final T-TAQ items. Although these comparisons revealed few meaningful differences—largely due to the fact that all respondents reported overall positive attitudes towards teamwork—one interesting finding resulted. With regard to the overall questionnaire, physicians and dentists combined had a significantly lower overall scale mean (4.07) than did nurses (4.11). However, both of these means represent positive attitudes towards teamwork overall. Leadership was the only construct for which a subgroup difference was identified, such that nurses had a significantly more positive view (4.53) of leadership than did physicians (4.23). However, despite this difference, both subgroups’ means represent highly positive attitudes towards aspects of teamwork related to leadership. Finally, no differences in overall attitudes towards teamwork were found between the military and civilian samples.

DISCUSSION
The T-TAQ provides a potentially useful tool for assessing individual attitudes related to the role of teamwork in the delivery of healthcare. The data reported here suggest that the tool is reliable. Preliminary analyses also indicate that the T-TAQ likely possesses discriminant validity, though additional research is required to confirm the true construct validity of the T-TAQ. Additional tests using confirmatory factor analysis, item response modelling or other advanced analysis techniques are warranted. Limiting our analyses to classical item statistics is a potential weakness of this study, though more advanced techniques require larger sample sizes to produce stable results.

From both a practical and research perspective, the T-TAQ can be used in a number of ways. First, it can be used to diagnose existing attitudes towards teamwork in a system, hospital, or unit. Based on the results, an organisation can determine if staff is ready for team training. Extremely negative results may indicate that other factors need to be resolved prior to implementing the team training initiative. However, little research exists on whether or not positive attitudes towards teamwork are required for team training to be successful in healthcare, and the relation between attitudes and behaviour is muddy at best. Therefore, future research should test this argument.

Second, because the T-TAQ focuses on core teamwork skills, we argue that it can be used to support quality improvement activities associated with teamwork beyond just the implementation of TeamSTEPPS. Although the T-TAQ is closely aligned with TeamSTEPPS, the T-TAQ could be used in congruence with other team training programmes that target the core skills, as well as in cases where researchers simply want to understand existing attitudes towards teamwork.

Third, consistent with Kraiger et al, the T-TAQ may be used to evaluate the effectiveness of team training by assessing one aspect of learning. When administered prior to and after...
training, data from the questionnaire can be used to assess changes in participant attitudes towards teamwork as a result of training. In this way, T-TAQ provides a resource to trainers by addressing Kirkpatrick’s Level 2 evaluation standard. However, our results showed that participant attitudes were positive without having received team training. Moreover, past research in other domains has shown that it is difficult to develop assessments that measure an individual’s knowledge of teamwork; the items tend to be too easy or highly correlated with general intelligence. Therefore, team training effectiveness in healthcare at Kirkpatrick’s Level 2 may only be defined by significant improvements in skills, because participants are likely to have positive attitudes and know what to do. Future research needs to test this argument.

Finally, the T-TAQ has significant potential to stimulate broader research on attitudes towards teamwork in healthcare and other domains. Of interest are those factors that contribute to the development of such positive or negative attitudes, how past experiences in teams affect attitudes towards teamwork, and whether or not positive attitudes are related to better team process in the delivery of care. These and other questions can only be pursued and answered when carefully constructed measurement tools are available.

**Funding** This work was supported under a contract to the Department of Defense as a subcontractor to Stanley Associates (Contract No W81XW11-06-F-0526).

**Competing interests** None.

**Ethics approval** Ethics approval was provided by the American Institutes for Research.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**REFERENCES**