How Do Scores on the Functional Status Examination (FSE) Correspond to Scores on the Glasgow Outcome Scale-Extended (GOSE)?

Lindsay D. Nelson,1,* Brooke E. Magnus,2 Nancy R. Temkin,3 Sureyya Dikmen,4 Geoffrey T. Manley,5 and Steve Balsis6

Abstract
This study was designed to determine how raw scores correspond between two alternative measures of functional recovery from traumatic brain injury (TBI), the Functional Status Examination (FSE) and the Glasgow Outcome Scale-Extended (GOSE). Using data from 357 persons with moderate-severe TBI who participated in a large clinical trial, we performed item response theory analysis to characterize the relationship between functional ability measured by the FSE and GOSE at 6 months post-injury. Results revealed that raw scores for the FSE and GOSE can be linked, and a table is provided to translate scores from one instrument to the other. For example, a FSE score of 7 (on its 0–21 scale, where higher scores reflect more impairment) is equivalent to a GOSE score of 6 (where GOSE is scaled on an 8-point scale, with higher scores reflecting less impairment). These results allow clinicians or researchers who have a score for a person on one instrument to cross-reference it to a score on the other instrument. Importantly, this enables researchers to combine data sets where some persons only completed the GOSE and some only the FSE. In addition, an investigator could save participant time by eliminating one instrument from a battery of tests, yet still retain a score on that instrument for each participant. More broadly, the findings help anchor scores from these two instruments to the broader continuum of injury-related functional limitations.

Keywords: functional limitations; Functional Status Examination; Glasgow Outcome Scale-Extended; item response theory; outcome measurement
Introduction
The Glasgow Outcome Scale-Extended (GOSE) and Functional Status Examination (FSE) are two well-recognized, leading, and validated interview-based measures of functional limitations after traumatic brain injury (TBI).\(^5\) A growing body of research indicates that these two instruments measure the same construct of injury-related functional limitations, implying one may be a reasonable substitute for the other.\(^5\) Each instrument has advantages, such as widespread recognition of the GOSE’s 1–8 scale and the ability of the FSE, which has a wider score range, to more precisely characterize individual differences in TBI recovery.\(^5\) Given that these two leading instruments are used in research studies and clinical trials, there is a need to understand how scores on the two correspond.

Knowing how scores on the FSE and GOSE correspond would enable several research goals. For example, being able to translate GOSE to FSE scores and vice versa would allow researchers to combine data sets that administered only one of these measures. Or, an investigator who had limited time with a research participant might decide to administer just one instrument, while retaining information about how that participant would have scored on the other instrument. In this way, understanding how two instruments correspond may help with ongoing goals of developing an efficient yet precise outcome measurement for clinical trials.\(^5\) Recently, TBI researchers have turned their attention to advanced psychometric techniques, including item response theory (IRT), which allows for the linking of scores between measures by placing both measures on a common scale (a latent continuum, in this case injury-related functional impairment). Leveraging our team’s recent work supporting the use of such models with the GOSE and FSE, we sought to link raw scores for each measure and create a table that researchers and clinicians can use to translate GOSE scores to FSE scores and vice versa. We also describe our approach to IRT score linking, which may be used to link other important instruments widely used in TBI research and clinical practice.

Methods
Participants
We performed secondary data analyses of patients with moderate-to-severe TBI who were enrolled in a randomized controlled trial of magnesium sulfate at Harborview Medical Center (Seattle, WA) between August 1998 and October 2004. This original study showed no positive effect of the magnesium sulfate intervention on outcomes, but the investigators found a negative effect of treatment in the lower-dosage group.\(^6\) For this original study, participants were required to have intracranial surgery within 8 h of injury, a post-resuscitation Glasgow Coma Scale (GCS)\(^1\) score of 3–12 or, if intubated, a GCS motor score of 1–5 without pharmacological paralysis. Patients who were under 14 years of age, unable to receive the study drug within 8 h of injury, who had serum creatinine concentrations greater than 177 \(\text{\mu mol/L}\), were pregnant, a prisoner, or who were living overseas were excluded from the analyses. Given the focus of the present study on 6-month measures of functional limitations, we analyzed data from the subset of participants \((n = 357)\) who had sufficient outcome data at 6 months to be included in IRT models. The selection of this subsample is described further in Nelson and colleagues.\(^8\)

Outcome assessment
As described in previous publications using this data set,\(^6,18,19\) participants completed diverse neuropsychological assessments at various points in the first 6 months post-injury. Relevant to the current study, the GOSE and FSE were collected at 6 months post-injury. The GOSE was collected using the recommended interview guidelines in place at that time.\(^2\) In particular, questions about independence in and outside the home, work/school functioning, social functioning, psychological and relationship functioning, and other functional limitations were included. Responses to these questions were used to place patients on an 8-point ordinal scale: 1 = Death, 2 = Vegetative State, 3 = Lower Severe Disability, 4 = Upper Severe Disability, 5 = Lower Moderate Disability, 6 = Upper Moderate Disability, 7 = Lower Good Recovery, and 8 = Upper Good Recovery. For this analysis, we retained cases with scores that ranged between 2 and 8 inclusive because these are the outcome ranges for which FSE scores are formally defined.

The FSE structured interview was conducted with patient participants and/or their significant other (close family member or friend). Past work showed a strong correlation (0.80–0.82) between patient and significant other ratings\(^6,18\) and no IRT differential item functioning by respondent.\(^8\) The study used the original version of the FSE interview, which contained 10 subsections: Personal Care; Mobility/Ambulation;
Mobility/Travel; Primary Activity (among work and school); Home Management; Leisure and Recreation; Social Integration; Cognitive & Behavioral Competency; Standard of Living; and Financial Independence. However, because the latter three domains were not retained in revisions of the instrument by its author, these domains were not included in the present analysis. Each domain on the FSE is rated on a 4-point ordinal scale, from 0 (same level of function as before injury) to 3 (complete or nearly complete reliance on others or inability to function in the domain), and these scores are summed to produce a total score reflecting the degree of injury-related functional impairment (range, 0–21 using seven domains, although persons who died were assigned an FSE score of 22 in the clinical trial). Both the FSE and GOSE interviews in this study included injury-related functional limitations regardless of whether their cause was TBI or other concurrent injuries.

Statistical analyses
Analyses for the present study used IRT parameters established in a previous publication from this sample. This model entered the seven FSE items and the GOSE overall score into a single model to establish the relationship between these items and the latent continuum of injury-related functional limitations. In summary, the data met the two necessary assumptions for conducting IRT analyses—that the data represent one underlying dimension and are locally independent. For example, the first/second eigenvalue ratio from an exploratory factor analysis was sufficiently large (18.0), and the fit of a one-factor confirmatory factor analytical model adequate, to consider the data sufficiently unidimensional. A unidimensional IRT model was estimated using IRTPRO version 4.2 to quantify the difficulty (b) and discrimination (a) of FSE and GOSE items. In IRT terms, difficulty reflects the level of disability severity (the latent continuum; called theta) indicated by the item; discrimination reflects the strength of the relationship between each item and the latent continuum. (Higher discrimination means that the item yields more precise estimates of one’s score on the latent continuum.)

As is described further in the Results, these item parameters were used to model the relationship between FSE and GOSE observed scores and theta scores and to perform single-group concurrent calibration (i.e., linking FSE to GSOE scores according to where they relate on the theta [latent disability] scale). We modeled our linking approach after one implemented in previous psychometric efforts.

Results
IRT model parameters quantified the relationship between FSE and GOSE scores and the latent continuum of compromised functional ability. This model was used to plot the test characteristic curve (i.e., relationship between test scores and latent impairment level) for the FSE and GOSE across the continuum of functional impairment (Fig. 1) and develop a table linking GOSE to FSE scores (see Table 1). Figure 1 shows that as FSE scores increase from 0 to 21 (see left axis), the test characteristic curve for the FSE rises across the continuum of compromised functional ability (horizontal axis). This monotonically increasing curve is consistent with the design of the FSE, which characterizes greater disability with higher scores. In contrast, the equivalent curve for the GOSE is monotonically decreasing (from 8 to 2) because the GOSE characterizes greater disability with lower scores (see right axis in Fig. 1).

This latent variable (theta), compromised functional ability, provides a common scale on which to link the FSE and GOSE. In other words, each point along the continuum between −3.00 and 3.00 standard deviations (SDs) corresponds both to a point along the FSE curve and a point along the GOSE curve. For example, note that a score of 7 on the FSE corresponds to a score of 6 on the GOSE, because each raw score corresponds to a value of −0.41 SDs of compromised functional ability. Figure 1 helps to illustrate the overall findings. In order to derive more precise/quantitative estimates of the relationship between the latent continuum, FSE, and GOSE scores, we created Table 1, which uses the IRT model parameters to translate the data reflected in Figure 1 into columns of values in Microsoft Excel (Microsoft Corporation, Redmond, WA), thereby establishing a value for the FSE and GOSE at every point along the theta continuum between −3.00 and 3.00, in increments of 0.01 SD units. Note that this produced estimated FSE and GOSE scores that were decimals.

With these values established, we traced the FSE scores from 0 to 21, breaking them into integer scores based on traditional rounding rules (e.g., estimated FSE 0.00–0.49 we coded as the integer score 0, estimated FSE between 0.50 and 1.49 coded as the integer score 1, all the way up to an FSE score of 21). These ranges...
of theta associated with each predicted FSE score are presented in Table 1. Then, GOSE scores were extracted from the point along the continuum closest to each raw FSE score (i.e., point closest to FSE score of 0, 1, 2, etc.; see Table 1). The resulting Table 1 provides a cross-walk of GOSE, FSE, and theta scores, which can be used by clinicians and researchers to determine how a patient’s raw score on one instrument, say the FSE, corresponds to a score on the GOSE.

Note that because FSE item-level scores are not formally defined for persons who are dead, and this IRT model relied on item-level data, this model included living persons only (i.e., GOSE ≥2). The range of FSE scores derived from the model is 0–21 (see Fig. 1) because the model was based on its seven items (each scaled 0–3, where 3 reflects that a person is either not participating in an outcome domain or is completely dependent on support to do so, which can occur because of severe impairment or being in a vegetative state). This scaling, in turn, formally restricts the model-estimated maximum FSE score to 21. However, from Table 1, it is apparent that an FSE score of 21 could reflect either a GOSE 2 or 3, an assignment that would require clinical information to differentiate. This is because persons with a 21 FSE score were rated a 3 (maximum impairment) on all items, which could occur because of being in a vegetative state or otherwise being severely globally unable to participate in major life domains independently. Further, although the FSE summed score is not formally defined for persons who died, based on past recommendations one could assign an FSE score of 22 to persons with GOSE 1 (dead) on conceptual grounds.16

Discussion
The GOSE and FSE are both respected measures of TBI-related functional limitations that have been used in clinical studies of TBI. Historically, formal guidance was not available to identify how scores on these instruments are related. The current study used a process of IRT linking to show how FSE and GOSE scores correspond to each other and in relation to the level of functional limitations reflected by different FSE and GOSE scores. We also described the method for IRT linking, which may provide a template for other attempts at linking measures important to TBI research. Further, we created a table that includes each score on the two measures and the corresponding score on the other test that indicates the same level of functional ability. As one would expect,
our findings demonstrate that most GOSE scores reflect ranges of multiple possible FSE scores, meaning that the GOSE is a coarser representation of function than the FSE.

Researchers and clinicians can use the score conversion table to better leverage existing data and better characterize TBI patients. For example, the findings anchor GOSE and FSE scores to the level(s) of the broader continuum of functional limitations reflected by those scores. Importantly, the findings could be used to combine data sets from studies regardless of whether they administered the FSE or GOSE. In light of proposals to use the FSE as a more granular outcome measure than the GOSE for clinical trials, the findings could also inform how adopting the FSE as a primary outcome measure in clinical trials would work, as compared to traditional outcome measurement strategies (e.g., to identify how differing GOSE cutoffs for “favorable outcome” translate in FSE terms).

The relatively simple, straightforward method of IRT linking described here may prove useful to other aspects of TBI research. For example, IRT could be used to develop score conversion tables for other widely used instruments of mild TBI (mTBI)-related symptoms, quality of life, and other constructs or even to map instruments from differing constructs (e.g., cognitive performance, symptoms, or, potentially, biomarkers) to each other by charting where they fall on the latent continuum of TBI-related disability. The method described in the current study could be used any time the same sample of persons completes multiple measures, as is common in extended neuropsychological assessments. Beyond the scope of this article, IRT linking is a deep subspecialty with methods available for more complex study designs.

It is important to note that FSE-GOSE translations enabled by Table 1 reflect their linkage along the latent continuum of functional impairment, with potentially different meaning than traditional GOSE scores. For example, a GOSE score of 5 indicates that the patient can be left alone at home for 24 h, whereas this is not guaranteed by an IRT-predicted GOSE score. On the other hand, it is possible that IRT-based scores are more precise or valid than the traditional GOSE score because of being characterized according to the latent continuum of functional limitations based on all interview responses provided. Additional research would be useful to explicate the similarities and distinctions between traditional observed scores on these instruments and the latent continuum they can be considered to reflect.

This GOSE-FSE score conversion table significantly advances an understanding of how the GOSE and FSE measure different parts of the continuum of functional limitations while also explicitly mapping how their scores compare to each other. Note that because our model used impairment ratings provided irrespectively of their source (TBI vs. peripheral injuries), we cannot verify that the score mapping would be equivalent for ratings of the specific impact of TBI on patients’ functioning. Similarly, this model may not translate to other subpopulations of TBI (e.g., mTBI) or time points beyond 6 months post-injury, given that these instruments and the resultant IRT model have not yet been demonstrated to be invariant over severity groups and time. However, with this type of approach, it would be theoretically possible to harmonize longitudinal data across these two instruments. In summary, this study demonstrates how sophisticated

| GOSE<sup>b</sup> | FSE | Theta |
|---------------|---|------|
| 8             | 0  | -1.63|
|               | 1  | -1.63 to 1.20 |
| 7             | 2  | -1.19 to -0.97 |
|               | 3  | -0.96 to -0.79 |
|               | 4  | -0.78 to -0.65 |
|               | 5  | -0.64 to -0.53 |
|               | 6  | -0.52 to -0.42 |
| 6             | 7  | -0.41 to -0.32 |
|               | 8  | -0.31 to -0.22 |
|               | 9  | -0.21 to -0.12 |
|               | 10 | -0.11 to -0.02 |
| 5             | 11 | -0.01 to 0.08 |
|               | 12 | 0.09 to 0.19 |
|               | 13 | 0.20 to 0.31 |
|               | 14 | 0.32 to 0.44 |
| 4             | 15 | 0.45 to 0.59 |
|               | 16 | 0.60 to 0.75 |
|               | 17 | 0.76 to 0.95 |
|               | 18 | 0.96 to 1.17 |
| 3             | 19 | 1.18 to 1.42 |
|               | 20 | 1.43 to 1.82 |
|               | 21 | 1.83 to 2.36 |
| 2             | 21 | ≥2.37 |

Theta reflects the level of the latent construct of functional limitations modeled by item response theory (IRT), in standard deviation units.

<sup>a</sup>Although FSE total scores (sum of item scores) are formally defined as ranging from 0 to 21, its authors have previously assigned persons who died (GOSE 1) an FSE score of 22.

<sup>b</sup>GOSE score placed next to the FSE score for which it matched most closely in the model.

FSE, Functional Status Examination; GOSE, Glasgow Outcome Scale-Extended.
psychometrics can bridge gaps in knowledge about alternative measures important to TBI research and will help fuel continued advancements in modern outcome measurement approaches for clinical and translational studies of TBI.

**Authors’ Contributions**
Dr. Nelson contributed to the conceptualization of the study, data analysis, interpretation of the findings, and drafting of the manuscript. Dr. Magnus contributed to the conceptualization of the study, interpretation of the findings, and critically reviewed and approved the manuscript. Dr. Temkin contributed to data analysis decisions and critically reviewed and approved the final manuscript. Dr. Dikmen critically reviewed and approved the final manuscript. Dr. Manley critically reviewed and approved the final manuscript. Dr. Balsis contributed to the conceptualization of the study, data analysis, interpretation of the findings, and drafting of the manuscript.

**Funding Information**
This secondary analysis project was funded by National Institutes of Neurological Disorders and Stroke (NINDS) grant no. R01 NS10856 (PI Nelson) and U.S. Department of Defense grant W81XWH-14-2-0176 (PI Manley). The clinical trial was funded by NINDS grant no. R01 NS19643.

**Author Disclosure Statement**
Dr. Nelson receives funding for unrelated research from the Centers for Disease Control and Prevention, Department of Defense, National Institutes of Health, National Football League, and Medical College Advancing a Healthier Wisconsin endowment. Dr. Dikmen is the author of the Functional Status Examination. Dr. Manley receives a research endowment. Dr. Dikmen is the author of the Functional Status Examination, Department of Defense, National Institutes of Health, National Football League, and Medical College Advancing a Healthier Wisconsin endowment.

**References**
1. McMillan, T., Wilson, L., Ponsford, J., Levin, H., Teasdale, G., and Bond, M. (2016). The Glasgow Outcome Scale—40 years of application and refinement. Nat. Rev. Neurol. 12, 477–485.
2. Wilson, J.T., Pettitgrew, L.E., and Teasdale, G.M. (1998). Structured interviews for the Glasgow Outcome Scale and the extended Glasgow Outcome Scale: guidelines for their use. J. Neurotrauma 15, 573–585.
3. Jennett, B., Snoek, J., Bond, M.R., and Brooks, N. (1981). Disability after severe head injury: observations on the use of the Glasgow Outcome Scale. J. Neurol. Neurosurg. Psychiatry 44, 285–293.
4. Dikmen, S, Machamer J, Manley GT, Yuh EL, Nelson LD, and Temkin NR. (2019). Functional Status Examination versus Glasgow Outcome Scale. J. Neurotrauma 36, 2423–2429.
5. Hudak, A.M., Caesar, R.R., Frol, A.B., Krueger, K., Harper, C.R., Temkin, N.R., Dikmen, S., Carille, M., Madden, C., and Diaz-Arrastia, R. (2005). Functional outcome scales in traumatic brain injury: a comparison of the Glasgow Outcome Scale (Extended) and the Functional Status Examination. J. Neurotrauma 22, 1319–1326.
6. Dikmen, S., Machamer, J., Miller, B., Doctor, J., and Temkin, N. (2001). Functional status examination: a new instrument for assessing outcome in traumatic brain injury. J. Neurotrauma 18, 127–140.
7. Nelson, L.D., Brett, B.L., Magnus, B.E., Balis, S., McCreight, M.A., Manley, G.T., Temkin, N., and Dikmen, S. (2020). Functional Status Examination yields higher measurement precision of functional limitations after traumatic injury than the Glasgow Outcome Scale-Extended: a preliminary study. J. Neurotrauma 37, 675–679.
8. Nelson, L.D., Magnus, B.E., Temkin, N., Dikmen, S., and Balis, S. (2021). Functional Status Examination yields higher measurement precision than the Glasgow Outcome Scale-Extended after moderate-to-severe traumatic brain injury. J. Neurotrauma 38, 3288–3294.
9. Nelson, L.D., Ranson, J., Ferguson, A.R., Giacino, J., Okonkwo, D.O., Valadka, A., Manley, G., and McCreight, M. (2017). Validating multidimensional outcome assessment using the TBI Common Data Elements: an analysis of the TRACK-TBI Pilot sample. J. Neurotrauma 34, 3158–3172.
10. Narayan, R.K., Michel, M.E., Ansell, B., Baethmann, A., Biegon, A., Bracken, M.B., Bullock, M.R., Choi, S.C., Clifton, G.L., Contant, C.F., Coplin, W.M., Dietrich, W.D., Ghajar, J., Grady, S.M., Grossman, R.G., Hall, E.D., Heftetrs, W., Hovda, D.A., Jallo, J., Katz, R.L., Knoller, N., Kochanek, P.M., Maas, A.J., Majde, J., Marion, D.W., Marmaro, A., Marshall, L.F., McIntosh, T.K., Miller, E., Mohberg, N., Muizelaar, J.P., Pitts, L.H., Quinn, P., Riesenfeld, G., Robertson, C.S., Strauss, K.J., Teasdale, G., Temkin, N., Tuma, R., Wade, C., Walker, M.D., Weinrich, M., Whyte, J., Wilberger, J., Young, A.B., and Yurkewicz, L. (2002). Clinical trials in head injury. J. Neurotrauma 19, 503–557.
11. Shukla, D., Devi, B.I., and Agrawal, A. (2011). Outcome measures for traumatic brain injury. Clin. Neurol. Neurosurg. 113, 435–441.
12. Bullock, M.R., Merchant, R.E., Choi, S.C., Gilman, C.B., Kreutzer, J.S., Marmaro, A., and Teasdale, G.M. (2002). Outcome measures for clinical trials in neuromodulation. Neurosurg. Focus 13, ECP1.
13. Wilson, J.T., Slieker, F.J., Legrand, V., Murray, G., Stocchetti, N., and Maas, A.I. (2007). Observer variation in the assessment of outcome in traumatic brain injury: Experience from a multicenter, international randomized clinical trial. Neurosurgery 61, 123–128; discussion, 128–129.
14. Maas, A.I., Braakman, R., Schouten, H.J.A., Minderhoud, J.M., and van Zomeren, A.H. (1983). Agreement between physicians on assessment of outcome following severe head injury. J. Neurosurg. 58, 321–325.
15. Maas, A.I., Roosenbeek, B., and Manley, G.T. (2010). Clinical trials in traumatic brain injury: past experience and current developments. Neurotherapeutics 7, 115–126.
16. Temkin, N.R., Anderson, G.D., Winn, H.R., Ellenbogen, R.G., Britz, G.W., Schuster, J., Lucas, T., Newell, D.W., Mansfield, P.N., Machamer, J.E., Barber, J., and Dikmen, S.S. (2007). Magnesium sulfate for neuroprotection after traumatic brain injury: a randomised controlled trial. Lancet Neurol. 6, 29–38.
17. Teasdale, G., Maas, A., Lecky, F., Manley, G., Stocchetti, N., and Murray, G. (2004). The Glasgow Coma Scale at 40 years: standing the test of time. Lancet Neurol. 13, 844–854.
18. Machamer, J., Temkin, N.R., Manley, G.T., and Dikmen, S. (2018). Functional Status Examination in patients with moderate-to-severe traumatic brain injuries. J. Neurotrauma 35, 1132–1137.
19. Machamer, J., Temkin, N., and Dikmen, S. (2013). Health-related quality of life in traumatic brain injury: is a proxy report necessary? J. Neurotrauma 30, 1845–1851.

20. Zahniser, E., Temkin, N.R., Machamer, J., Barber, J., Manley, G.T., Markowitz, A.J., and Dikmen, S.S.; TRACK-TBI Investigators. (2019). The Functional Status Examination in mild traumatic brain injury: a TRACK-TBI sub-study. Arch. Clin. Neuropsychol. 34, 1165–1174.

21. Slocum-Gori, S.L. and Zumbo, B.D. (2011). Assessing the unidimensionality of psychological scales: using multiple criteria from factor analysis. Soc. Indic. Res. 102, 443–461.

22. Hu, L., and Bentler, P.M. (1999). Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. Struct. Equ. Modeling 6, 1–55.

23. Balsis, S., Benge, J.F., Lowe, D.A., Geraci, L., and Doody, R.S. (2015). How do scores on the ADAS-Cog, MMSE, and CDR-SOB correspond? Clin. Neuropsychol. 29, 1002–1009.

24. Hudak, A.M., Caesar, R.R., Frol, A.B., Krueger, K., Harper, C.R., Temkin, N.R., Dikmen, S.S., Carllle, M., Madden, C., and Diaz-Arrastia, R. (2005). Functional outcome scales in traumatic brain injury: a comparison of the Glasgow Outcome Scale (Extended) and the Functional Status Examination. J. Neurotrauma 22, 1319–1326.

Cite this article as: Nelson, LD, Magnus, BE, Temkin, NR, Dikmen, S, Manley, GT, and Balsis, S (2022) How do scores on the Functional Status Examination (FSE) correspond to scores on the Glasgow Outcome Scale-Extended (GOSE)? Neurotrauma Reports 3:1, 122–128, DOI: 10.1089/neur.2021.0057.

Abbreviations Used

- FSE = Functional Status Examination
- GCS = Glasgow Coma Scale
- GOSE = Glasgow Outcome Scale-Extended
- IRT = Item response theory
- mTBI = mild TBI
- SDs = standard deviations
- TBI = traumatic brain injury

Publish in Neurotrauma Reports

- Immediate, unrestricted online access
- Rigorous peer review
- Compliance with open access mandates
- Authors retain copyright
- Highly indexed
- Targeted email marketing

liebertpub.com/neur