P1.5 AN ASSESSMENT OF THE VARIABILITY IN OPERATIONAL ASSIGNMENT OF F-SCALE DAMAGE

Jared L. Guyer *
NOAA/National Weather Service Hastings, Nebraska

Todd J. Shea
NOAA/National Weather Service La Crosse, Wisconsin

1. INTRODUCTION

On average, there are approximately 1000 tornadoes in the continental United States each year (Storm Data). Because of human tendencies, varying degrees of experience, and a myriad of additional factors, it is widely accepted that there is a high degree of subjectivity involved in assigning F-scale (Fujita 1971) ratings to tornadoes.

In an attempt to quantify the variability of F-scale assignment, the authors conducted an internal National Weather Service (NWS) F-scale rating exercise. This survey expands on a previous rating study by Edwards and Harmon (1999) and a concurrent paper by Edwards (2003). The authors felt it would be important to sample NWS personnel, since F-scale ratings are almost exclusively assigned by NWS meteorologists.

The intent of this exercise was not a "right vs. wrong" test or a "second guess" of the original F-scale rating, but instead, an attempt to measure the potential variability of F-scale assignment in a quantitative sense given a varying assortment of tornado damage pictures.

2. METHODOLOGY

An e-mail was distributed internally within the NWS to all Warning Coordination Meteorologists (WCMs), Meteorologists-In-Charge (MICs), and Science and Operations Officers (SOOs) to introduce and explain the rating exercise. There are approximately 121 field offices within the NWS, with one WCM, one MIC, and one SOO in each office. Information about this exercise was redirected within some of the NWS offices, hence the limited participation by NWS meteorologists other than WCMs, MICs, or SOOs.

WCMs have the primary responsibility for post-storm damage assessment in the NWS, but MICs and SOOs also conduct surveys and play a role in the process. In general, other meteorologists within NWS field offices perform damage surveys less often.

* *Corresponding author address:* Jared L. Guyer, National Weather Service, 6365 Osborne Drive West, Hastings, NE 68901; e-mail: Jared.Guyer@noaa.gov

Within the e-mail introduction to this exercise, participants were directed to a website (http://www.crh.noaa.gov/gid/fscale/) featuring 13 sets of damage pictures.

Instructions stated that all pictures were of tornado induced damage. Given each set of pictures and limited information, the participants were asked to provide their best estimation of the appropriate F-scale rating, in addition to an accompanying justification for their rating.

Aside from the F-scale rating and justification, other demographic information was asked to potentially determine comparisons between geographic regions or experience level. Exercise participants submitted their responses electronically and these results were assembled into a database.

Figure 1. Damage pictures (#1 and #8 respectively) from F-scale web exercise – http://www.crh.noaa.gov/gid/fscale/

3. RESULTS

A total of 56 NWS meteorologists participated in the F-scale rating exercise, which is approximately 15 percent of the original invitation to all NWS WCMs, MICs, and SOOs. Forty-four percent of the exercise respondents were WCMs (Table 2).

Of the respondents, the average amount of NWS tenure was 15 years, with a minimum of 1 year, and a maximum of 39 years. The respondents had participated in an average number of 37 damage surveys (tornadic and non-tornadic) in their NWS

career, with an average of 14 damage surveys involving tornadoes. When asked the maximum F-scale they had experienced, the average response was an F-scale rating of 3.3 (additional information Table 3).

NWS Position	*Percentage of Overall Participants*
WCMs	44%
SOOs	20%
MICs	13%
Other Meteorologists	23%

Table 2. F-scale exercise participants by NWS position.

Maximum F-scale Rating Experienced	*Percentage of Overall Participants*
F0	7%
F1	2%
F2	13%
F3	20%
F4	27%
F5	21%
N/A	10%

Table 3. Maximum F-Scale rating participants had experienced in their career. Average maximum F-scale rating was 3.3.

	Average F-scale Rating	*Standard Deviation*	*Mode*	*F-scale Range*
Picture 1	1.6	0.49	2	1 – 2
Picture 2	1.5	0.66	1	0 – 3
Picture 3	2.5	1.20	2	0 – 5
Picture 4	2.8	0.62	3	1 – 4
Picture 5	1.7	0.54	2	0 – 3
Picture 6	1.1	0.51	1	0 – 2
Picture 7	1.6	0.91	1	0 – 4
Picture 8	2.7	0.67	3	1 – 4
Picture 9	3.3	0.66	3	2 – 4
Picture 10	2.1	1.10	2	0 – 5
Picture 11	1.4	0.76	1	0 – 4
Picture 12	2.9	0.67	3	1 – 4
Picture 13	4.0	0.70	4	2 – 5

Table 4. Overall average F-scale rating, standard deviation, mode, and range for each set of damage pictures.

In general, a broad variability was observed in the F-scale ratings assigned in this exercise. Ranges of 4 or more F-scale ratings were common with each set of pictures (Table 4). For the 13 exercise pictures, standard deviations ranged from 0.49 to 1.20, with an overall "average" standard deviation for the exercise of 0.73 (calculated from Table 4).

A trend of slightly lower F-scale ratings and smaller standard deviations was observed of NWS Southern Region meteorologists in comparison to NWS Central and Western Region meteorologists (Eastern Region excluded because of a distinctly smaller response size); however, evidence of tendencies based upon NWS position or NWS region (Tables 5 and 6) were largely inconclusive given a relatively limited data set. A clearer distinction was evident when responses were delineated by overall damage experience, as those with more experience (greater than 20 surveys) responded with a lower average F-scale rating and a smaller standard deviation (Table 7).

Upon completion of the F-scale rating exercise, participants had an opportunity to offer additional comments and feedback concerning the F-scale and this exercise (Table 8).

4. SUMMARY AND CONCLUSIONS

The response ranges and standard deviations observed in the study were similar to that of the previous exercise by Edwards and Harmon (1999) and the concurrent paper by Edwards (2003). How these survey measurements apply to real-world occurrences is certainly debatable given the inherent limitations (e.g., "static" pictures and limited information) of such an exercise. However, it could be argued that the standard deviations (Table 4) from the exercise appear to fall into reasonable ranges of what might naturally occur given differences of opinion and interpretation on F-scale ratings.

With little exception, the respondents did not have formalized engineering training. Given this, and the broad variability of F-scale observed in the exercise, the authors propose formalized training, with the assistance of wind and structural engineers, be further incorporated into the NWS training curriculum. Furthermore, per supporting comments specified in the exercise, the authors suggest a formalized listing of damage assessment references, training materials, and standards be distributed to all NWS field offices. While the F-scale will always remain subjective by its very nature, it is theorized that enhanced training may partially mitigate this variance.

More specifically, the authors propose the development of web-based or CD-ROM training modules (e.g., COMET or VISIT modules) with the assistance of wind and structural engineers to simulate actual damage assessment scenarios. The authors feel this could ultimately be a valuable asset to training in damage assessment practices. It is speculated with ever evolving technology, that a virtual reality type of instructional tool could eventually best mimic a true-to-life damage assessment setting.

5. ACKNOWLEDGEMENTS

The authors would like to thank the NWS personnel who participated in this exercise. Their time and comments were greatly appreciated. A special thanks to Roger Edwards for his insight into this project. Constructive reviews by Glenn Lussky, Al Pietrycha, Cynthia Fay, and Rick Ewald greatly contributed to this manuscript.

6. REFERENCES

Fujita, T.T., 1971: Proposed characterization of tornadoes and hurricanes by area and intensity. SMRP Res. Paper 91, Univ. of Chicago, 42 pp.

Edwards, R. and D.G. Harmon, 1999: Lubbock F-Scale Exercise on Spencer, SD Tornado Damage. Conducted at NWS/Texas Tech Severe Weather Conference, Lubbock TX.

Edwards, R., 2003: Rating Tornado Damage: An Exercise in Subjectivity. Preprints, Symposium on F-Scale and Severe-Weather Damage Assessment, Long Beach, CA, this volume.

Storm Data. National Oceanic and Atmospheric Administration, National Climatic Data Center, Asheville, North Carolina.

	Average F-scale Rating by NWS Position ***WCM** (24) / **SOO** (11) / **MIC** (7) / **Other** (14)*	***Standard Deviation by NWS Position*** ***WCM / SOO / MIC / Other***	***Mode by NWS Position*** ***WCM / SOO / MIC / Other***
Picture 1	1.5 / 1.8 / 1.9 / 1.5	0.51 / 0.40 / 0.38 / 0.52	2 / 2 / 2 / 1
Picture 2	1.5 / 1.3 / 2.0 / 1.5	0.66 / 0.66 / 0.82 / 0.52	1 / 1 / 2 / 1
Picture 3	2.4 / 2.7 / 2.9 / 2.5	0.92 / 1.27 / 1.21 / 1.61	3 / 2 / 2 / 2
Picture 4	2.5 / 2.8 / 2.7 / 3.1	0.72 / 0.40 / 0.49 / 0.49	3 / 3 / 3 / 3
Picture 5	1.7 / 1.6 / 1.6 / 1.8	0.47 / 0.81 / 0.53 / 0.44	2 / 2 / 2 / 2
Picture 6	1.0 / 1.2 / 1.4 / 1.1	0.46 / 0.40 / 0.53 / 0.64	1 / 1 / 1 / 1
Picture 7	1.4 / 1.9 / 2.0 / 1.5	0.78 / 0.94 / 1.29 / 0.85	1 / 2 / 2 / 1
Picture 8	2.6 / 2.7 / 2.7 / 2.7	0.72 / 0.47 / 0.76 / 0.73	3 / 3 / 2 / 3
Picture 9	3.3 / 3.5 / 3.1 / 3.2	0.69 / 0.52 / 0.69 / 0.73	3 / 3 / 3 / 3
Picture 10	2.0 / 2.2 / 2.1 / 2.0	1.18 / 0.98 / 0.90 / 1.22	1 / 3 / 3 / 2
Picture 11	1.3 / 1.5 / 1.6 / 1.4	0.81 / 0.82 / 0.79 / 0.65	1 / 1 / 1 / 1
Picture 12	2.7 / 2.8 / 3.1 / 3.1	0.76 / 0.60 / 0.69 / 0.47	3 / 3 / 3 / 3
Picture 13	3.9 / 4.0 / 4.3 / 3.9	0.76 / 0.77 / 0.52 / 0.62	4 / 4 / 4 / 4

Table 5. Average F-scale rating, standard deviation, and mode by NWS position.

	Average F-scale Rating by NWS Region ***CR** (26) / **SR** (15) / **WR** (9)*	***Standard Deviation by NWS Region*** ***CR / SR / WR***	***Mode by NWS Region*** ***CR / SR / WR***
Picture 1	1.7 / 1.6 / 1.6	0.49 / 0.51 / 0.53	2 / 2 / 2
Picture 2	1.6 / 1.3 / 1.6	0.64 / 0.46 / 1.01	1 / 1 / 1
Picture 3	2.8 / 2.2 / 2.9	1.24 / 0.77 / 1.45	3 / 2 / 2
Picture 4	2.8 / 2.6 / 2.9	0.65 / 0.63 / 0.60	3 / 3 / 3
Picture 5	1.8 / 1.7 / 1.3	0.49 / 0.46 / 0.71	2 / 2 / 1
Picture 6	1.2 / 0.9 / 1.3	0.51 / 0.52 / 0.50	1 / 1 / 1
Picture 7	1.7 / 1.5 / 1.6	1.09 / 0.83 / 0.73	2 / 1 / 1
Picture 8	2.6 / 2.7 / 3.0	0.76 / 0.49 / 0.71	3 / 3 / 3
Picture 9	3.6 / 3.0 / 3.3	0.58 / 0.53 / 0.71	4 / 3 / 3
Picture 10	2.2 / 1.3 / 3.1	0.85 / 0.96 / 1.05	2 / 1 / 2
Picture 11	1.5 / 1.4 / 1.4	0.71 / 0.99 / 0.53	1 / 1 / 1
Picture 12	3.0 / 2.7 / 2.9	0.63 / 0.82 / 0.60	3 / 3 / 3
Picture 13	4.0 / 4.0 / 4.1	0.72 / 0.68 / 0.83	4 / 4 / 4

Table 6. Average F-scale rating, standard deviation, and mode by NWS region (*CR = Central Region, SR = Southern region, WR = Western Region*). Note: Eastern Region excluded from comparison because of limited sample size (3).

	Average F-scale Rating by Experience ***≤ 20 Surveys** (33) / **> 20 Surveys** (18)*	***Standard Deviation by Experience*** ***≤ 20 / > 20***	***Mode by Experience*** ***≤ 20 / > 20***
Picture 1	1.6 / 1.6	0.49 / 0.50	2 / 2
Picture 2	1.6 / 1.4	0.75 / 0.50	1 / 1
Picture 3	2.7 / 2.4	1.31 / 0.92	2 / 2
Picture 4	2.8 / 2.7	0.66 / 0.59	3 / 3
Picture 5	1.6 / 1.7	0.60 / 0.47	2 / 2
Picture 6	1.2 / 1.1	0.55 / 0.42	1 / 1
Picture 7	1.8 / 1.3	0.94 / 0.91	1 / 1
Picture 8	2.8 / 2.4	0.70 / 0.61	3 / 2
Picture 9	3.3 / 3.4	0.68 / 0.50	3 / 3
Picture 10	2.2 / 2.0	1.19 / 0.91	2 / 2
Picture 11	1.4 / 1.3	0.76 / 0.75	1 / 1
Picture 12	2.8 / 2.9	0.70 / 0.64	3 / 3
Picture 13	4.0 / 3.9	0.68 / 0.73	4 / 4

Table 7. Average F-scale rating, standard deviation, and mode by number of overall damage surveys participated in career.

This type of thing would be nice for all WCMs as a pre-Spring drill. Some review and explanation would be good for each if used for training.
…It would be much appreciated to have a highly expanded description of each F-scale rating for better clarity in making F-scale decisions in the future.
Good cross section of damage. Like the real world, many of these are tough calls, even if we had the benefit of being able to walk around and get a closer look. So many variables to consider!
It is really hard to tell, based on pictures alone. Also, had to make some assumptions on the type of construction and distances to surrounding trees, buildings, etc. that could not be accurately determined from the photos.
…Much needed topic for clarification and consistency with storm surveys.
Very difficult to make a fair assessment on just the picture without truly being at the site.
I suspect your study will verify the difficulty of consistently assigning damage an appropriate F-scale rating. Hope the variation you get in ratings is truly due to differences in opinion, and not affected by lack of data.
I'd like to see the survey results and "correct" answers. I wonder if I have a bias one way or another.

Table 8. Selected comments by F-scale exercise participants.