JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, January 2015, Vol. 60, No. S1
doi: 10.1111/1556-4029.12616
Available online at: onlinelibrary.wiley.com

TECHNICAL NOTE

CRIMINALISTICS

Gregory E. Gorbett,[1,2] *M.S.; Sarah M. Morris,*[2] *M.S.; Brian J. Meacham,*[1] *P.E., Ph.D.; and Christopher B. Wood,*[1,3] *P.E., J.D.*

A New Method for the Characterization of the Degree of Fire Damage to Gypsum Wallboard for Use in Fire Investigations

ABSTRACT: A new method to characterize the degree of fire damage to gypsum wallboard is introduced, implemented, and tested to determine the efficacy of its application among novices. The method was evaluated by comparing degree of fire damage assessments of novices with and without the method. Thirty-nine "novice" raters assessed damage to a gypsum wallboard surface, completing 66 ratings, first without the method, and then again using the method. The inter-rater reliability was evaluated for ratings of damage without and with the method. For novice fire investigators rating degree of damage without the aid of the method, ICC(1,2) = 0.277 with 95% CI (0.211, 0.365), and with the method, ICC(2,1) = 0.593 with 95% CI (0.509, 0.684). Results indicate that the raters were more reliable in their analysis of the degree of fire damage when using the method, which support the use of standardized processes to decrease the variability in data collection and interpretation.

KEYWORDS: forensic science, fire investigation, fire patterns, inter-rater reliability, gypsum wallboard, fire damage, calcination

Fire investigators use visible observations of fire damage (damage indicators) as their principal means of determining a fire's area of origin (1–6). Some fire investigators have historically used, and in some cases continue to use, these damage indicators as a mechanism by which decisions are made to explain the physical evidence presented without necessarily having a good understanding of the underlying causative fire dynamics. This may have contributed to the promulgation of several myths within the profession (1). Within the past decade, increased scrutiny of the fire investigation profession has been made public through the United States media reporting on several miscarriages of justice due to myths and the lack of standardized procedures used in the past (1,2). Sadly, in many areas of the United States, these procedures and prolific use of the myths continue. Over the past decade, momentum has increased within the fire investigation profession to move away from this indexed list of explanations and move toward the use of discrete damage indicators in an attempt to understand the dynamics of how the fire developed and caused the resulting damage (1).

Fire and arson investigation is possibly one of the most complicated facets of the forensic sciences, due to the end result being frequently path independent. Identical or nearly identical damage indicators may result from different fire scenarios leading to a multitude of plausible hypotheses, meaning that multiple paths of the fire could result in similar damage. As such, it would be expected that there are processes and limitations established to assist with the identification of varying degrees of damage remaining after a fire and the link that degree of damage has to the origin of the fire. Currently, however, no process or methodology exists that permits an objective or uniform identification of varying levels of fire damage. Many fire investigation reports, textbooks, and standards inconsistently report degrees of damage, using a wide range of vague modifiers, such as greater, lesser, heavy, light, major, moderate, minor, severe, and large, in an attempt to distinguish between levels of damage that they observe and are trying to convey (1–6). The absence of a formal process combined with the use of vague modifiers when reporting on data that serves as the principal support for an investigator's conclusions results in several major problems. These include unpredictable conclusions, inter-rater and intrarater reliability issues, and validity issues (7,8). Such factors can be seen as shortcomings to admissibility standards of scientific evidence as laid out in *Daubert v. Merell Dow Pharmaceuticals* (9) in the United States and *R v. Mohan* (10) in Canada.

To address these concerns, this study involves the development of a structured methodology, which can be used to guide identification and characterization of damage indicators. Development of the method is aligned with recommendations from the National Academy of Science review of forensic sciences in the United States, to establish standard terminology and undertake research that address issues of reliability and validity in forensic science (11). Both of these recommendations are fundamental to assist the fire investigation profession. Other forensic science and engineering disciplines (12,13) have benefited significantly

[1]Department of Fire Protection Engineering, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609.

[2]Department of Fire Protection and Paramedicine Sciences, Fire, Arson, and Explosion Investigation Program, Eastern Kentucky University, 521 Lancaster Avenue, Richmond, KY 40475.

[3]FireLink, LLC, 1501 Main Street, Suite 17, Tewksbury, MA 01876.

Received 18 June 2013; and in revised form 29 Oct. 2013; accepted 27 Jan. 2014.

from developing clear parameters for identification purposes and standardizing their lexicon, which in turn has permitted a deeper evaluation of their respective forensic science and allowed for integration of advancements in technology.

This paper discusses the development of a degree of fire damage method to characterize gypsum wallboard damage and the evaluation of the inter-rater reliability without and with the application of this method. Gypsum wallboard consists of a core of gypsum (calcium sulfate dihydrate) sandwiched between two thick paper facers (14). Gypsum wallboard has a predictable response to heat and its uniformity in production allow it to be used as a reliable indicator of heat exposure for postfire analysis (1,17). There are several effects that may occur to gypsum wallboard when exposed to heat and fire conditions, including color changes, soot deposition, texture changes, charred paper, consumed paper, and clean burn (soot is not observed postfire and appears to have been consumed). Additionally, when gypsum wallboard is exposed to heat, it will undergo a dehydration of chemically bound water, known as calcination, leaving a fragile material in its place (14,17).

Methods

Within this article, four distinct activities are discussed: (i) the development of a degree of fire damage scale, (ii) methodology to apply the scale, (iii) a study of the method as applied by novices, and (iv) statistical analysis to evaluate the method's effectiveness.

Measures

Typically, fire investigators look at the face of all surface linings after a fire and make visible determinations of the varying degree of fire damage (DOFD). Gypsum wallboard-lined walls and ceilings are one of the most common lining materials utilized in residential and commercial construction. As such, gypsum wallboard will serve as the most beneficial material to begin the development of a method to objectively characterize the DOFD and will be the focus of this study.

As a first step, a DOFD scale was developed as a ranking system to reflect the varying degrees of visible fire damage to gypsum wallboard based on its response to heat exposure and visible damage indicators (VDI). The VDI and their respective varying degrees of damage were compiled from the literature, drawing from the many texts and research studies that detail the impact of heating to gypsum wallboard (14,17). Next, a scale ranging from 0 to 6 was developed for assigning a DOFD, with 0 indicating no visible damage and 6 indicating complete consumption. Each level within the scale was based on a set of VDIs outlined by the literature review. These VDIs were detailed within each level to characterize the DOFD. The VDIs included color and texture differences. Selected images of the VDI for each level were also provided with the DOFD scale to serve as examples to assist with the analysis (Table 1 and Fig. 1).

A method of characterizing fire damage observed along gypsum wallboard-lined surfaces was developed from combining the ranking scale, example images, description of damage indicators, and instructions on how to apply the method (Table 1). To identify varying DOFD along larger surfaces, it is necessary to increase the resolution through the use of a grid system. As such, the user is instructed to establish an appropriate grid size for the surface being evaluated. The user would then evaluate

TABLE 1—*Method of characterizing degree of fire damage along Gypsum wallboard-lined surfaces.*

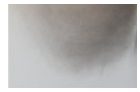
Degree of Fire Damage (DOFD)	Visible Damage Indicator Description (VDI)	Selected Images of Visible Damage Indicators (corresponding to Fig. 1)
0	*No visible damage:* these areas are noted by their original surface color (white if unpainted; painted surface color)	
1	*Soot deposited on surface:* these areas are noted by discoloration of the original surface color; but the facing paper is still present	
2	*Discoloration of facing paper and loss of paint:* these are locations of the gypsum wallboard surface that have discolored due to thermal effects, the paper can be brownish, light black, or dark black in color; Or variations in color depending on original paint color	
3	*Paper is beginning to peel, bubble, and flake*: the paper has been penetrated and the gypsum wallboard is exposed	
4	*The paper has been consumed*: these areas are typically gray or white in color	
5	*Clean burn*: near complete consumption of paper and soot accompanied by a white/bluish color	
6	*Complete consumption:* complete loss of integrity and mass of the gypsum wallboard	
N/A	Damage cannot be determined due to a suppression or unknown causes	

FIG. 1—*Photograph from which selected images of visible damage indicators was chosen for use with the DOFD Method.*

each grid space and characterize the damage within that space. The method further instructed the user to use the example images and damage indicators to characterize the damage observed. Additional instructions were provided to clarify those potential areas of difficulty in ranking. These instructions indicated that the user should be conservative and select the degree of damage with the lower value in the event that the grid space had two varying degrees of damage of equivalent areas (i.e., if a grid space is half soot covered and half no damage, then the user should select no damage for this grid space). Furthermore, instructions were provided that if the grid space includes a seam in the drywall that had been covered with drywall tape and finishing compound, then the participant should determine the most prevalent degree of damage for the gypsum wallboard and ignore the effects of the tape.

Study Design and Sample

To test the reliability of the proposed method, participants (novices) were asked to complete a characterization exercise of a color photograph of a fire damaged gypsum wallboard-lined wall first without the method and then again with the method.

Volunteers were asked to participate as novices applying the method in the study. The participants included 39 undergraduate students in their first course in fire investigation with no formal training or practical experience. Although this was not a random sample, the participants were reasonably representative of typical novices. A single color photograph of a wall damaged from exposure to known fire conditions was chosen for this series of observational tests. An alphanumeric grid was superimposed on the photograph; the columns of the grid were labeled A-K beginning at the left edge of the image, while the rows were labeled 1–6 beginning at the top of the image (Fig. 1). Each of the 66 individual grid spaces encompassed an area of *c.* 0.14 m^2 (0.375 × 0.375 m) (Fig. 2).

Procedures

Each of the 39 novice participants was supplied with the photograph and 66-grid overlay. First, they were asked to rank the most prevalent damage for each grid space based on a scale from 0–6, with 0 indicating no damage and 6 indicating complete consumption. The novice participants performed this first analysis without any methodology and were expected to assign varying degrees of fire damage on their own (as investigators typically do in the field). Next, the participants were provided the same photograph and 66-grid overlay and were asked to provide a rating of 0–6 of the most prevalent degree of damage for each grid space using the degree of fire damage method (Table 1). They were instructed to carefully read through the method and use it as a reference when identifying damage within a grid space of the photograph.

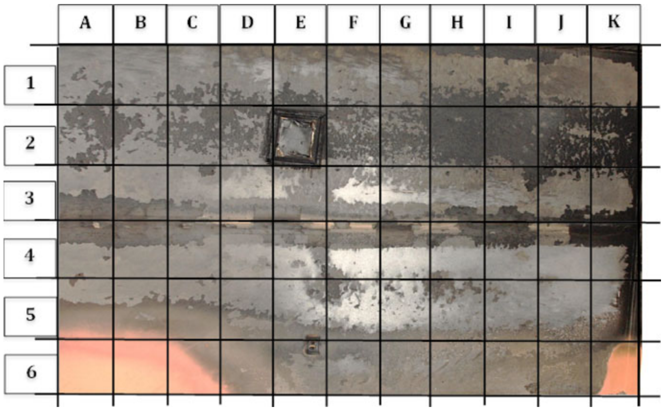


FIG. 2—*Photograph of damaged wall with superimposed grid.*

Ratings from participants were collected electronically using Qualtrics survey software (15). This platform provided the participants with a simple method to record the damage rating for each of the 66 grid spaces by utilizing a dropdown selection menu containing only the values 0–6. The participants were able to return to grid spaces throughout each study and correct errant values. However, once the study was submitted, they were no longer able to access their answers. The participants were not permitted to talk to each other as they performed the study. Due to the relatively large number of cells being evaluated, participant fatigue was a concern. An attention verification question was asked in the middle of the survey to ensure that participants were actively engaged in selecting answers and not haphazardly choosing values. Three participants and their results were excluded for failing the attention validation test.

Data Analysis

To assess the reliability of the DOFD method among participants, the intraclass correlation coefficient (ICC) was calculated for the 39 participants. The ICC is a descriptive statistic that quantitatively estimates rater reliability on a scale from 0 to 1 with a higher value indicating stronger agreement between raters. Specifically, ICC (2,1) was selected as each participant rated each of the 66 grid spaces, and absolute agreement was chosen to account for systematic error due to the relatively small sample size (16). Strength of agreement was interpreted according to the following scale to maintain nomenclature consistent with other reliability measures (18):

ICC(2,1)	Strength of Agreement
<0.40	Poor
0.40–0.75	Fair to good
>0.75	Excellent

Finally, a paired *t*-test was conducted to determine whether there was a difference in mean overall damage ratings for novices without and with the method. ICCs were calculated using SPSS version 19 for Windows (19); all *t*-tests were performed using SAS version 9.2 (20). A significance level of $\alpha = 0.05$ was used throughout.

Results

For novice fire investigators rating degree of damage without the aid of the DOFD method, ICC(2,1) = 0.277 with 95% CI (0.211, 0.365). This relatively small ICC value indicates high variability in ratings and poor agreement among participants. For novice ratings making use of the DOFD method, ICC (2,1) = 0.593 with 95% CI (0.509, 0.684), indicating fair to good agreement among participants.

The mean damage rating for novices without the DOFD method was 3.32 (SD = 0.54), while the mean rating with the DOFD method was 3.57 (SD = 0.34). It is interesting to note that the mean value of damage significantly increased with the use of the DOFD method ($t = 3.52$, $p = 0.001$), despite the instructions to be conservative in ranking damage (Fig. 3).

Discussion

As the ICC increases from 0.277 to 0.593 for novices without and with the DOFD method, respectively, the agreement

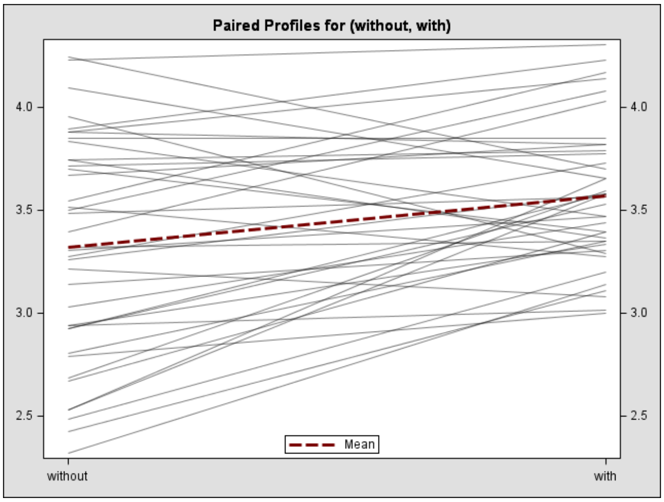


FIG. 3—*Change in overall "novice" rating without method (left) to with DOFD method (right), illustrating decrease in variance.*

among raters is increasing and the variability in their ratings is decreasing. This decrease in variability of the novice's ratings with the method illustrates the reliability of the DOFD method.

The simple DOFD method presented in this study has been shown to decrease the variability among novices and increases the reliability in ranking fire damage to gypsum wallboard.

Even though this study has shown a reduction in variability in the degree of fire damage among novices, it is interesting to determine whether these results are similar to that of expert practitioners. Due to time constraints, a limited convenience sample of four expert fire investigators was used to rate the degree of fire damage using the method. The expert ratings had almost no variability (mean overall damage rating for the experts was 3.62; SD = 0.04). It is interesting to note that novices had an overall damage rating of 3.32 without the method, which increased to 3.57 with the method. This trend, despite the limited expert sample, may indicate that novices when using the method are rating damage similarly to the expert practitioners. Further testing will need to be conducted with expert practitioners to further evaluate this trend.

The DOFD method will need to be further tested against a variety of fire damaged gypsum wallboard-lined walls with various paint schemes to further evaluate its reliability and consistency.

Conclusion

A new method is presented for characterizing the degree of fire damage to gypsum wallboard-lined surfaces. Thirty-nine independent "novice" raters performed a visual analysis of damage to a wall surface, completing 66 ratings first without the degree of fire damage method and second, repeated rating with the DOFD method. The inter-rater reliability was evaluated for ratings of damage without and with the method. The results indicate that the novice raters were more reliable in their analysis of the degree of fire damage to gypsum wallboard when using the DOFD method. These results support the use of standardized processes to decrease the variability in data collection and interpretation.

References

1. National Fire Protection Association. NFPA 921: the guide for fire and explosion investigations. Quincy, MA: NFPA, 2011.
2. DeHaan J, Icove J. Kirk's fire investigation, 7th edn. Upper Saddle River, NJ: Brady/Pearson, 2011.
3. Madrzykowski D, Fleischmann C. Fire pattern repeatability: a study in uncertainty. J Test Eval 2012;40(1):11.
4. Shanley J, Kennedy P, Ward J. Report of the United States Fire Administration program for the study of fire patterns. Washington, DC: Federal Emergency Management Agency, 1997; Report No.: FA 178, 7/97.
5. Milke J, Hill S. Full-scale room fire experiments. Gaithersburg, MD: National Institute of Science and Technology, 1996; Report No.: NIST GCR-96-703.
6. Rethoret H. Fire investigations. Toronto, ON, Canada: Recording and Statistical Corp, Ltd., 1945.
7. Gann D. Trial by fire: did Texas execute an innocent man? The New Yorker 2009 September 7; Sect. A:3.
8. Arizona's Pioneer Hotel Fire Re-examined [videorecording]. New York, NY: 60-minutes CBS, 2013 March 31.
9. Daubert v. Merrell Dow Pharmaceuticals, Inc. 509 U.S. 579 (1993).
10. R v. Mohan. S.C.R. 9. Supreme Court of Canada (1994).
11. National Academy of Science. Strengthening forensic science in the United States: a path forward. Washington, DC: National Academies Press, 2009.
12. National Oceanic and Atmospheric Administration. A guide to F-scale damage assessment. Silver Spring, MD: U.S. Department of Commerce, 2007.
13. Mealy C, Gottuk D. Full-scale validation tests of a forensic methodology to determine smoke alarm response. Fire Technol 2010;47:275–89.
14. Schroeder R. Post-fire analysis of construction materials [dissertation]. Berkeley, CA: University of California Berkeley, 1999.
15. Qualtrics Labs, Inc. Qualtrics [computer program]. Version 40829. Provo, UT: Qualtrics Labs, Inc., 2005.
16. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86(2):420–8.
17. Mann DC, Putaansuu ND. Studies of the dehydration/calcination of Gypsum wallboard. Proceedings of Fire and Materials; 2009 January 26–28; San Francisco, CA. London, U.K.: Interscience, 2009;827–35.
18. Fleiss JL. Statistical methods for rates and proportions. New York, NY: John Wiley, 1981.
19. SPSS Inc. SPSS [computer program]. Windows Version 19.0. Chicago, IL: SPSS Inc., 2014.
20. SAS Institute Inc. SAS [computer program]. Version 9.2 ed. Cary, NC: SAS Institute Inc., 2005.

Additional information and reprint requests:
Gregory E. Gorbett, M.S.
Eastern Kentucky University
250A Stratton Building
Richmond, KY 40475
E-mail: greg.gorbett@eku.edu