

2016

Categorizing Fetal Heart Rate Variability With and Without Visual Aids

Amanda J. Ashdown
Old Dominion University, aashd001@odu.edu

Mark W. Scerbo
Old Dominion University, mscerbo@odu.edu

Lee A. Belfore II
Old Dominion University, lbelfore@odu.edu

Stephen S. Davis
Eastern Virginia Medical School

Alfred Z. Abuhamad
Eastern Virginia Medical School

Follow this and additional works at: https://digitalcommons.odu.edu/psychology_fac_pubs



Part of the [Bioimaging and Biomedical Optics Commons](#), [Cardiology Commons](#), [Industrial and Organizational Psychology Commons](#), and the [Pediatrics Commons](#)

Original Publication Citation

Ashdown, A. J., Scerbo, M. W., Belfore, L. A., II, Davis, S. S., & Abuhamad, A. Z. (2016). Categorizing fetal heart rate variability with and without visual aids. *AJP Rep*, *06*(04), e359-e366. <https://doi.org/10.1055/s-0036-1593605>

This Report is brought to you for free and open access by the Psychology at ODU Digital Commons. It has been accepted for inclusion in Psychology Faculty Publications by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

Categorizing Fetal Heart Rate Variability with and without Visual Aids

Amanda J. Ashdown, MS¹ Mark W. Scerbo, PhD, FHFES¹ Lee A. Belfore II, PhD² Stephen S. Davis, MD³
 Alfred Z. Abuhamad, MD³

¹ Department of Psychology, Old Dominion University, Norfolk, Virginia

² Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, Virginia

³ Department of Obstetrics and Gynecology, Eastern Virginia Medical School, Norfolk, Virginia

Address for correspondence Amanda J. Ashdown, MS, Department of Psychology, Old Dominion University, 5115 Hampton Boulevard, Norfolk, VA 23529 (e-mail: aashd001@odu.edu).

Am J Perinatol Rep 2016;6:e359–e366.

Abstract

Objective This study examined the ability of clinicians to correctly categorize images of fetal heart rate (FHR) variability with and without the use of exemplars.

Study Design A sample of 33 labor and delivery clinicians inspected static FHR images and categorized them into one of four categories defined by the National Institute of Child Health and Human Development (NICHD) based on the amount of variability within absent, minimal, moderate, or marked ranges. Participants took part in three conditions: two in which they used exemplars representing FHR variability near the center or near the boundaries of each range, and a third control condition with no exemplars. The data gathered from clinicians were compared with those from a previous study using novices.

Results Clinicians correctly categorized more images when the FHR variability fell near the center rather than the boundaries of each range, $F(1,32) = 71.69$, $p < 0.001$, partial $\eta^2 = 0.69$. They also correctly categorized more images when exemplars were available, $F(2,64) = 5.44$, $p = 0.007$, partial $\eta^2 = 0.15$. Compared with the novices, the clinicians were more accurate and quicker in their category judgments, but this difference was limited to the condition without exemplars.

Conclusion The results suggest that categorizing FHR variability is more difficult when the examples fall near the boundaries of each NICHD-defined range. Thus, clinicians could benefit from training with visual aids to improve judgments about FHR variability and potentially enhance safety in labor and delivery.

Keywords

- ▶ categorization
- ▶ cues
- ▶ clinical judgment
- ▶ exemplars
- ▶ fetal heart rate variability

Analysis of the fetal heart rate (FHR) is important for monitoring fetal wellbeing.^{1–3} Guidelines for interpreting electronic fetal monitoring (EFM) tracings are provided by the National Institute of Child Health and Human Development (NICHD) which defines FHR variability based on its amplitude.⁴ There are four categories of variability: (1) absent, FHR variability with an undetectable amplitude range; (2) minimal, FHR variability with an amplitude range of 5 or fewer bpm, (3)

moderate, FHR variability with an amplitude range of 6 to 25 bpm, and (4) marked, FHR variability with an amplitude range greater than 25 bpm.

Although the NICHD classification system provides a framework for FHR assessment, there is still potential for clinicians to misclassify FHR variability. This can result in the inappropriate operative intervention (i.e., cesarean procedures) and even increase the cost of obstetrics and malpractice insurance.^{2,3}

received
 July 12, 2016
 accepted after revision
 September 2, 2016

DOI <http://dx.doi.org/10.1055/s-0036-1593605>.
 ISSN 2157-6998.

Copyright © 2016 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA.
 Tel: +1(212) 584-4662.

License terms



Research has revealed poor agreement among clinicians when classifying FHR tracings⁵ and that judging FHR variability is difficult for suspicious and pathological tracings.⁶ To appreciate the challenge associated with judging FHR variability requires some understanding of how individuals categorize information.

Researchers who study category formation and judgment argue that individuals generate a category prototype through their experience with exemplars. Each new stimulus is compared with a mental prototype and classification judgments are based on degrees of similarity. Stimuli that are more like the prototype are easier to classify. Stimuli that deviate further from the prototype become more difficult to categorize.^{7–9} Stimuli falling near a boundary between two categories are more ambiguous and can impede the judgment process. Research also shows that when categorizing ambiguous stimuli, cues, and visual aids can facilitate performance because they direct attention to salient information.^{10–12} According to structural alignment theory, when individuals make comparisons features and relations within one stimulus are systematically matched (i.e., aligned) to features and relations in another stimulus, thereby isolating relevant information.^{13,14} Thus, an alignable (high-similarity) comparison standard can improve classification performance by making differences among stimuli more salient.¹¹

Several researchers recently studied the category judgment process for discriminating among levels of FHR variability.¹⁵ They asked a sample of university students to classify images representing the four NICHD categories of FHR variability: absent, minimal, moderate, and marked.⁴ They examined the classification of FHR variability using two types of images. Prototypical images depicted FHR variability at the midpoint of the four variability ranges and boundary images depicted FHR variability near the boundaries of the four categories. Also, they examined performance with and without visual aids incorporating alignable cues. Their results showed that participants categorized more images correctly when cues were available. Further, more images depicting variability near the boundaries of each range were correctly categorized with the presence of prototypical cues (i.e., exemplars representing FHR variability near the center of each range) than with exemplars representing FHR variability falling near the boundaries of each range.

The results from this initial study suggest that providing visual aids that incorporate alignable cues can improve judgments about FHR variability, especially when the images fall near the boundaries of the categories. However, the data were obtained from a naïve undergraduate population. Clinicians, however, spend many years classifying FHR tracings. Because clinicians have extensive experience examining FHR variability, they may have internalized prototypical representations of the four NICHD-defined categories of FHR variability. Thus, it is possible that visual aids might be less beneficial for experienced clinicians. On the other hand, alignable cues may also help clinicians categorize FHR variability, particularly when examples fall near the category boundaries. Thus, the goals of this study were to examine the ability of labor and delivery clinicians to categorize FHR variability with and without visual aids and to determine whether experts benefit from using visual aids when examples of FHR variability fall near the boundaries of each category.

Methods

Participants

Participants consisted of 33 clinicians working in labor and delivery (21 nurses, 10 residents, 1 faculty physician, and 1 midwife) from Eastern Virginia Medical School and Sentara Norfolk General Hospital with experience ranging from 3 to 25 years. All nurses, the midwife, and nine residents were females, and the faculty physician and one resident were males.

Fetal Heart Rate Display

The images used in this study were generated with a maternal-FHR simulator created by Belfore et al.¹⁶ Static images were taken from sections of dynamic tracings. An example image is displayed in ►Fig. 1.

Although FHR variability varies along a continuum, the NICHD segmented the continuum into the four categories as shown in ►Table 1. The prototypical images were created by using a value near the middle of each category of FHR variability (e.g., the prototypical image of the moderate category had a variability of 15 bpm). The boundary images were created by using values near the boundaries of each FHR variability category (e.g., a boundary image of the moderate category had a variability of 7 bpm near the lower end of the category or 20 bpm near the higher end of the category). The exemplars provided as visual aids were created by taking snapshots of the values of FHR variability used in the prototypical and boundary images and labeling them with the appropriate category of variability. They were placed side-by-side and under the FHR tracing. The visual exemplars are shown in ►Fig. 2.

Procedure

The participants were asked to participate by email and the experiment took place in a conference room at the Sentara Hospital in Norfolk. The residents and nurses consented before the start of the study.

Participants were shown sample images of each variability type to become familiar with the FHR images produced by the simulator and given the opportunity to practice categorizing five examples of each cue condition with feedback. The participants were assigned at random to begin the experiment with one of three sets of cue conditions (see below).

After the practice, the participants were seated at a computer to begin the experiment. They were presented with three blocks of images: 180 example images presented with exemplars (90 with prototypical cues and 90 with boundary cues) and 90 images without exemplars. Block order was counterbalanced across participants. There were 10 images created for each category prototype (i.e., absent, minimal, moderate, and marked) and 10 images near the boundaries of the categories. Each image was available for inspection until the participant made a classification decision. Participants responded by pressing a single key on the computer keyboard. Specific keys on the computer keyboard (A, V, M, and L) were labeled with the words “absent, minimal, moderate, or marked” for their convenience. Once

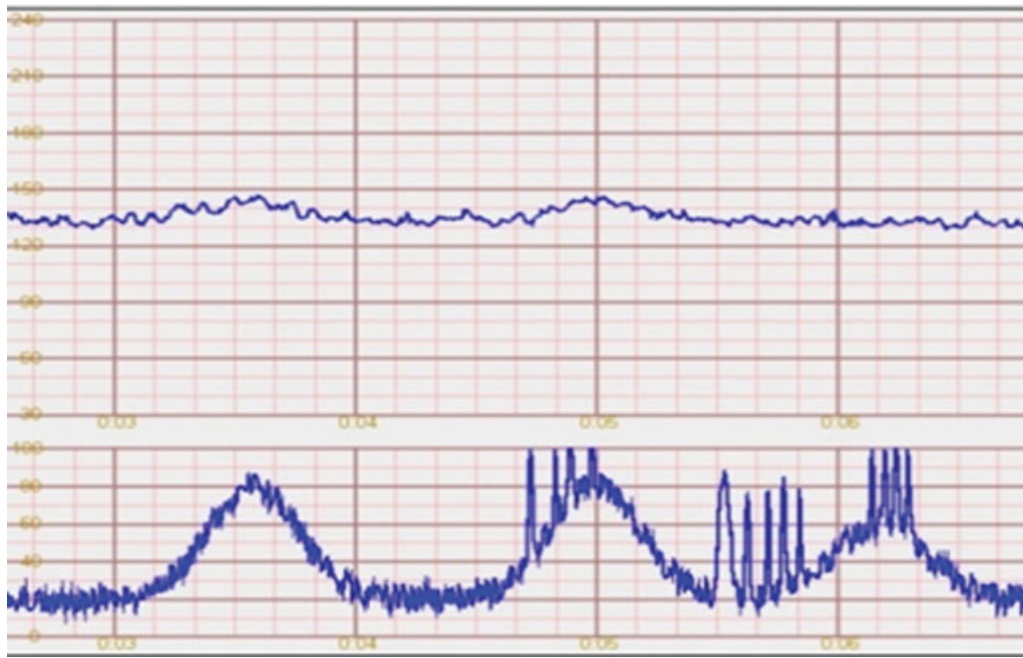


Fig. 1 Image of moderate FHR variability. The FHR is shown in the top portion of the display while the maternal contractions are shown in the bottom portion of the window. FHR, fetal heart rate.

participants made a response, the next example immediately was presented. Participants were given an opportunity to take a 5-minute break after each block. The computer informed them that the session had ended after they have completed both sets of images.

Design

The results were analyzed using a 3 (visual aid condition) × 2 (image type) analysis of variance (ANOVA) for a within-subjects design. The within-subjects variables were the presence of cues and the prototypicality of examples. The dependent measures were accuracy (i.e., the total number of examples each participant categorized correctly) and response times. A Sidak Bonferroni test was used for comparisons among means.

Results

Accuracy

Accuracy was analyzed by obtaining the number of correctly categorized examples in each condition divided by the total number of example images each condition contained, or pro-

Table 1 Standard variability ranges and the values selected for the prototypical and boundary bpm variabilities

Variability	Prototypical	Boundary
Absent 0 bpm	0 bpm	
Minimal > 0 and ≤ 5 bpm	3 bpm	1 and 4 bpm
Moderate 6–25 bpm	15 bpm	7 and 20 bpm
Marked > 25 bpm	35 bpm	30 bpm

Abbreviation: bpm, beats per minute.

portion of correct responses. A significant effect for image type was observed, $F(1,32) = 71.69, p < 0.001$, partial $\eta^2 = 0.69$. The participants correctly categorized more prototypical images (mean [M] = 0.91, standard deviation [SD] = 0.10) than boundary images (M = 0.75, SD = 0.11). A significant effect for the visual aid conditions was also observed, $F(2,64) = 5.44, p = 0.007$, partial $\eta^2 = 0.15$. A Sidak Bonferroni test revealed that the participants correctly categorized more images in the prototypical (M = 0.84, SD = 0.11) and boundary visual aid conditions (M = 0.84, SD = 0.10) compared with the no visual aid condition (M = 0.81, SD = 0.10).

In addition, a significant interaction was also observed for image type and visual aid condition, $F(2,64) = 3.72, p = 0.03$, partial $\eta^2 = 0.10$ (see **Table 2**). A Sidak Bonferroni test showed that for the prototypical images, participants correctly categorized significantly more images with the prototypical and boundary visual aids compared with the no visual aid condition. The posthoc test, however, showed no significant differences between the visual aid conditions for the boundary images.

Response Times

A significant main effect for image type was observed, $F(1,32) = 10.27, p = 0.003$, partial $\eta^2 = 0.24$. Participants took significantly longer to respond to the boundary images (M = 2.46, SD = 1.00) compared with the prototypical images (M = 2.10, SD = 0.97). No other effects were significant.

Comparison of Undergraduates and Clinicians

The data obtained from university undergraduates reported by Ashdown et al¹⁵ were compared with the data from the clinicians and analyzed using a 2 (experience) × 3 (cue

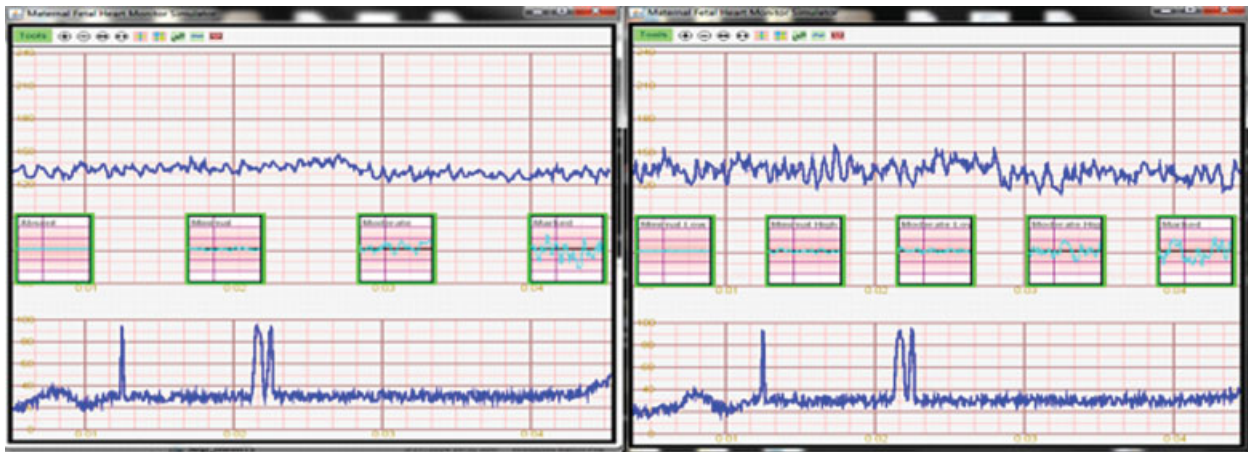


Fig. 2 The left image shows moderate FHR variability with prototypical visual aids placed below and the image on the right shows marked variability with boundary visual aids. FHR, fetal heart rate.

condition) × 2 (image type) ANOVA for a mixed design, with the cue conditions and the image type (prototypical or boundary) as the within-subjects factors and the experience level as the between-subjects factor. Descriptive statistics for both sets of data are shown in ► **Tables 2** and **3**.

Image Type

A significant effect for image type was observed, $F(1,72) = 186.6, p < 0.001, \text{partial } \eta^2 = 0.72$. Both clinicians and novices correctly categorized more prototypical images ($M = 0.90, SD = 0.10$) than boundary images ($M = 0.73, SD = 0.11$). A significant interaction was also observed for image type and cue condition, $F(1.82, 131.1) = 9.48, p < 0.001,$

$\text{partial } \eta^2 = 0.12$. A plot of the interaction is shown in ► **Fig. 3**. A test of simple effects showed that for the prototypical images, participants correctly categorized significantly more images in the boundary cue condition ($M = 0.94, SD = 0.08$) compared with both the prototypical ($M = 0.90, SD = 0.09$) and no-cue conditions ($M = 0.85, SD = 0.12$), and participants correctly categorized more images in the prototypical cue condition compared with the no-cue condition. The test of simple effects also showed that for the boundary images, participants correctly categorized more images in the prototypical ($M = 0.74, SD = 0.12$) and boundary cue condition ($M = 0.73, SD = 0.10$) compared with the no-cue condition ($M = 0.70, SD = 0.12$). Moreover, the

Table 2 Means and standard deviations for proportion of correct responses for clinicians

	Image type	Cue condition	Mean	Standard deviation
Clinicians	Prototypical image	Prototypical cue	0.92	0.10
		Boundary cue	0.93	0.10
		No cue	0.87	0.10
	Boundary image	Prototypical cue	0.76	0.11
		Boundary cue	0.75	0.09
		No cue	0.74	0.11

Table 3 Means and standard deviations for proportion of correct responses for novices

	Image type	Cue condition	Mean	Standard deviation
Novices	Prototypical image	Prototypical cue	0.88	0.10
		Boundary cue	0.94	0.07
		No cue	0.82	0.13
	Boundary image	Prototypical cue	0.73	0.12
		Boundary cue	0.71	0.10
		No cue	0.67	0.12

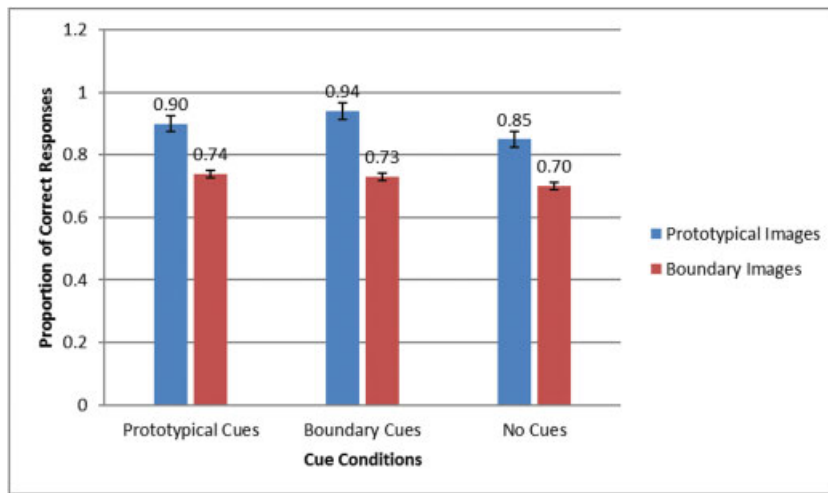


Fig. 3 Mean proportion of correct responses for image type as a function of each cue condition.

difference between prototypical and boundary images was significant in all three cue conditions, with fewer examples categorized correctly for the boundary images.

Cue Conditions

A significant effect for cue conditions was observed, $F(2,144) = 25.24, p < 0.001, \text{partial } \eta^2 = 0.26$. Participants correctly categorized more images in the prototypical ($M = 0.82, SD = 0.11$) and boundary ($M = 0.83, SD = 0.09$) cue conditions compared with the no-cue condition ($M = 0.78, SD = 0.12$).

Experience

A significant main effect for experience was observed, $F(1,72) = 4.9, p = 0.03, \text{partial } \eta^2 = 0.06$. Overall, the clinicians correctly categorized more images ($M = 0.83, SD = 0.10$) compared with the novices ($M = 0.79, SD = 0.11$); however, there was also a significant interaction between experience and cue condition, $F(2,144) = 3.36, p = 0.04, \text{partial } \eta^2 = 0.05$. A plot of the interaction is shown in **Fig. 4**. A test of simple effects showed that the clinicians

($M = 0.81, SD = 0.10$) correctly categorized significantly more images in the no-cue condition compared with the undergrads ($M = 0.74, SD = 0.12$). The clinicians also categorized more images in the boundary ($M = 0.84, SD = 0.10$) cue condition compared with the no-cue condition ($M = 0.81, SD = 0.10$). The novices correctly categorized more images in the prototypical ($M = 0.81, SD = 0.11$) and boundary ($M = 0.83, SD = 0.08$) cue conditions compared with the no-cue condition ($M = 0.74, SD = 0.12$). There were no significant differences between clinicians and novices in the prototypical and boundary cue conditions ($p > 0.05$).

Response Time

A significant main effect for experience was observed for response times, $F(1,72) = 5.78, p = 0.01, \text{partial } \eta^2 = 0.10$, in which novices took significantly longer to respond to images ($M = 2.89, SD = 1.62$) compared with clinicians ($M = 2.28, SD = 0.99$). A significant main effect for image type was also observed, $F(1,72) = 26.52, p < 0.001, \text{partial } \eta^2 = 0.30$. Both clinicians and novices took significantly

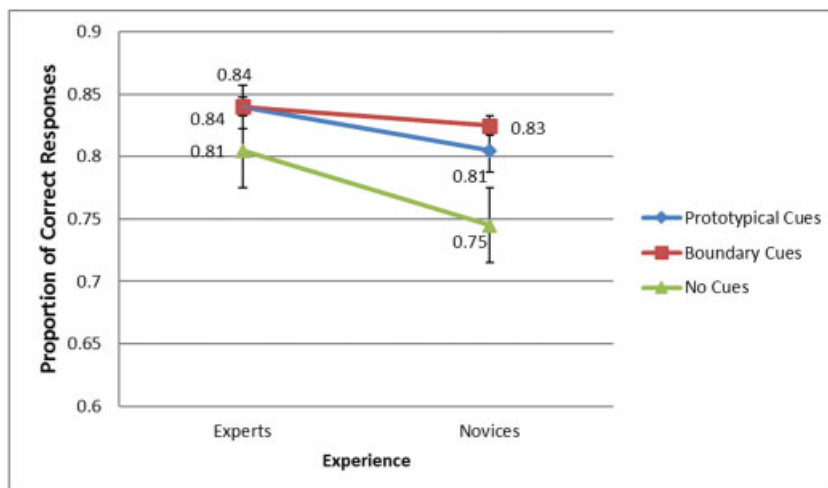


Fig. 4 Mean proportion of correct responses for each cue condition as a function of experience.

longer to respond to the boundary images ($M = 2.83$, $SD = 1.58$) compared with the prototypical images ($M = 2.34$, $SD = 1.23$). A significant interaction for image type and cue condition was observed, $F(1.84, 132.27) = 4.24$, $p = 0.02$, partial $\eta^2 = 0.06$. Regarding prototypical images, a test of simple effects showed that participants took longer to respond with the boundary cues ($M = 2.47$, $SD = 1.24$) than with the prototypical cues ($M = 2.16$, $SD = 1.00$). As for the boundary images, participants took longer to respond with the boundary cues ($M = 3.07$, $SD = 1.74$) than they did without cues ($M = 2.66$, $SD = 1.55$).

Discussion

The present study examined how well individuals could discriminate between the four categories of FHR variability defined by the NICHD: absent, minimal, moderate, and marked. Clinicians working in labor and delivery took part in three conditions: no cues, prototypical cues, and boundary cues. The prototypical cues represented the middle of each of the four FHR variability categories and the boundary cues represented the boundaries of each of the four FHR variability categories.

Image Type

The first goal was to assess whether the classification of FHR variability is more difficult as the examples deviate further from the prototype for each category. The results show that clinicians correctly categorized significantly more prototypical images compared with the boundary images, which supports other findings and suggests that examples of stimuli that are similar to their prototypes are easier to categorize.^{8,9,17} The participants also took longer to respond to the boundary images as compared with the prototypical images. These results support numerous studies that have found that examples of stimuli that resemble their prototypes are responded to more quickly.⁷⁻⁹

Cue versus No Cues

The second goal was to examine categorization performance when provided with exemplars. Previous research suggests that cues are able to direct attention to relevant information and help a person respond faster and make fewer errors when detecting stimuli in visual tasks.^{10-12,18-23} Clinicians in this study categorized more FHR variability images correctly when they had visual aids. In addition, for the prototypical images, the clinicians correctly categorized more images with the prototypical and boundary visual aids compared with the no visual aid condition. However, there were no significant differences between the visual aid conditions for the boundary images suggesting that categorizing examples of FHR variability that fall near the boundaries of each range is still a difficult task. The clinicians were less able to generalize to boundary images. This result suggests that when the images are farther away from the prototype within each category and closer to the boundaries, clinicians may be less able to rely on memory for making category decisions.

Comparison of Clinicians and Novices

The data obtained from novices in a previous study were compared with those of the clinicians. The results revealed that the clinicians correctly categorized more FHR images than the novices in all conditions, but statistically the advantage was limited to the no-cue condition. Differences in experience also depended on conditions. For the prototypical images, clinicians and novices correctly categorized more images in the boundary cue condition compared with the prototypical cue condition and both cue conditions compared with the no-cue condition. For the boundary images, all participants correctly categorized more images in the prototypical and boundary cue condition compared with the no-cue condition. Thus, it appears that clinicians and novices alike were able to benefit from the alignable cues. Again, alignable cues allow individuals to more easily compare information in examples by making relevant features more salient and improving category judgments.^{13,14}

The clinicians had more difficulty categorizing boundary images compared with prototypical images in all three conditions. At the outset, it was expected that experienced clinicians would be able to focus on the relevant details in the images and correctly categorize more boundary images in the no-cue and prototypical-cue condition, because previous research suggests that learning with exemplary diversity facilitates categorization of novel examples that deviate more from the category prototype.^{8,9,24-27} The clinicians had the advantage of years of learning from multiple examples; thus, they should have been able to organize the FHR variability categories around mental prototypes acquired through their experience and outperform the novices at categorizing boundary images. However, the results suggest that categorization of the FHR images may require more perceptual processing, guided by how similar each example is to the learned prototype of each category, making it difficult to categorize boundary images when needing to rely on memory. Thus, the task may be limited to the quality of one's perceptual discrimination.

Collectively, these results support previous research in which cues improved performance when categorizing maternal-FHR images,^{19,28} and suggest that the categorization of the FHR variability could be accomplished by pattern recognition.^{27,29} Both clinicians and novices were better able to categorize the examples by matching them to the alignable cues. In fact, when the cues were present, the clinicians had no appreciable advantage over the novices. Moreover, the results showed that when given the prototypical cues, the novices performed as well as the experts categorizing FHR variability without the cues.

Limitations

The main goal of this study was to examine how well individuals discriminate between categories of FHR variability when viewing static images of tracings. It should be noted that the FHR images in this study were created with a simulator and were not exact representations of real tracings. However, this approach was used because the primary focus was on characteristics of FHR variability and therefore it was imperative to specify precisely the variability present in the images.

This paradigm used here differs from standard clinical practice in another way. When maternity patients are monitored by EFM, the clinicians inspect dynamic tracings of the FHR. Therefore, the task of categorizing static images of FHR variability did not resemble the actual dynamic task clinicians perform on a regular basis in labor and delivery units because the participants had a limited view (5 minutes of a tracing), whereas, in hospitals, clinicians have an opportunity to inspect hours' worth of data. By narrowing and restricting the view, participants were not given the opportunity to use the additional information available in a longer FHR tracing. The clinicians may have been at a disadvantage because the context-detecting changes in the FHR variability were limited. Another potential limitation was that the experts differed in their level of experience. Future studies should look at how the training differs between residents and nurses because nurses are often the first to inspect the FHR tracings before calling the physician for further examination.

Theoretical and Clinical Implications

The results of the present study suggest that categorizing FHR variability is a difficult task because the boundaries of the NICHD categories are not perceptually distinct. Clinicians categorized significantly fewer boundary images compared with prototypical images and only correctly categorized approximately 81% of the FHR images when no cues were available. This suggests that the clinicians may have an internalized prototype of each variability category, but it may have limited utility when the examples fall near the border of each category, where categorization is most difficult. Some FHR images may be too difficult for clinicians to rely solely on memory for making category decisions. Thus, even experts could benefit from additional cues to more reliably categorize FHR variability.

Conclusion

In a clinical setting, the failure to distinguish among different types of FHR variability may impact the care of the patient. Clinicians must be able to judge the degree of variability to determine to which category of the three-tier FHR interpretation system the tracing belongs. If the variability is categorized incorrectly then there is a risk of misdiagnosis and unnecessary intervention. The present results showed that when no cues were available, clinicians were able to make their decisions using an internalized prototype of each FHR variability category acquired over their years of experience. However, categorizing examples of FHR variability near the boundaries of each range was still a difficult task. The findings suggest that when images deviate more from the prototype of each category and are closer to the boundaries, the task becomes more perceptual, making it difficult to rely on memory. When provided with cues, the clinicians may be able to make their judgments by pattern matching.

For unaided images, clinicians could draw upon their extensive experience with FHR tracings and be more adept at making their categorization decisions than novices. However, the presence of cues had an important impact on

performance. The advantage that clinicians enjoyed over novices in the no-cue condition was eliminated when cues were available. Thus, the availability of cues enabled novice university undergraduates to make categorization decisions comparable to those of clinicians.

Therefore, the presence of a visual cue can improve clinicians' abilities to properly categorize FHR variability when the tracings are prototypical examples of the categories, and may provide a benefit when the FHR variability is not prototypical. Because the visual aids benefitted everyone, it would be advantageous to examine whether training with the visual aids could improve performance. The current study suggests that the clinicians might benefit from training with alignable cues on categorizing FHR variability, which could improve later clinical decision making when monitoring FHR tracings to enhance the safety in labor and delivery.

References

- Freeman RK. Problems with intrapartum fetal heart rate monitoring interpretation and patient management. *Obstet Gynecol* 2002; 100(4):813–826
- Menihan CA, Zottoli EK. *Electronic Fetal Monitoring: Concepts and Applications*. Philadelphia, PA: Lippincott; 2001
- Sweha A, Hacker TW, Nuovo J. Interpretation of the electronic fetal heart rate during labor. *Am Fam Physician* 1999;59(9):2487–2500
- Macones GA, Hankins GDV, Spong CY, Hauth J, Moore T. The 2008 National Institute of Child Health and Human Development workshop report on electronic fetal monitoring: update on definitions, interpretation, and research guidelines. *Obstet Gynecol* 2008;112(3):661–666
- Figueras F, Albela S, Bonino S, et al. Visual analysis of antepartum fetal heart rate tracings: inter- and intra-observer agreement and impact of knowledge of neonatal outcome. *J Perinat Med* 2005;33(3):241–245
- Ayres-de-Campos D, Bernardes J, Costa-Pereira A, Pereira-Leite L. Inconsistencies in classification by experts of cardiotocograms and subsequent clinical decision. *Br J Obstet Gynaecol* 1999;106(12): 1307–1310
- Ashby FG, Boynton G, Lee WW. Categorization response time with multidimensional stimuli. *Percept Psychophys* 1994;55(1):11–27
- Das-Smaal EA, De Swart JH. Variation within categories. *Acta Psychol (Amst)* 1984;57:165–192
- Das-Smaal EA, De Swart JH. Effects of contrasting category, conjoint frequency and typicality on categorization. *Acta Psychol (Amst)* 1986;62:15–40
- Kok EM, de Bruin AB, Robben SG, van Merriënboer JJ. Learning radiological appearances of diseases: does comparison help? *Learn Instr* 2013;23:90–97
- Kurtz KJ, Gentner D. Detecting anomalous features in complex stimuli: the role of structured comparison. *J Exp Psychol Appl* 2013;19(3):219–232
- Wickens CD, Hollands JG. *Engineering Psychology and Human Performance*. 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall; 2000
- Gentner D, Gunn V. Structural alignment facilitates the noticing of differences. *Mem Cognit* 2001;29(4):565–577
- Markman AB, Gentner D. The effects of alignability on memory. *Psychol Sci* 1997; 8(5):363–367
- Ashdown AJ, Scerbo MW, Belfore LA II, Abuhamad AZ, Davis SS. Categorizing fetal heart rate variability with and without visual aids. In: *Proceedings of the 59th Annual Meeting of Human Factors & Ergonomics Society*; October 26–30, 2015; Los Angeles, CA
- Belfore LA II, Scerbo MW, Anderson BL. A fetal heart rate monitor simulator. Virginia Beach, VA: ModSim World; 2007

- 17 Vandierendonck A. Concept learning, memory and transfer of dot pattern categories. *Acta Psychol (Amst)* 1984;55(1):71–88
- 18 Chaney FB, Teel KS. Improving inspector performance through training and visual aids. *J Appl Psychol* 1967;51(4):311–315
- 19 Hall LM, Hannon DJ, Leisk GG, Wolfberg AJ, House MD. Measurement of fetal heart rate variability on an electronic monitor using a prototype electronic ruler. *Am J Perinatol* 2012;29(6):409–413
- 20 Loft S, Smith RE, Remington RW. Minimizing the disruptive effects of prospective memory in simulated air traffic control. *J Exp Psychol Appl* 2013;19(3):254–265
- 21 Gunn DV, Warm JS, Nelson WT, Bolia RS, Schumsky DA, Corcoran KJ. Target acquisition with UAVs: vigilance displays and advanced cuing interfaces. *Hum Factors* 2005;47(3):488–497
- 22 Posner MI, Nissen MJ, Ogden WC. Attended and unattended processing modes: The role of set for spatial location. In: Pick H, Saltzman E, eds. *Modes of Perceiving and Processing Information*. Hillsdale, NJ: Lawrence Erlbaum; 1978:137–157
- 23 Posner MI, Snyder CR, Davidson BJ. Attention and the detection of signals. *J Exp Psychol* 1980;109(2):160–174
- 24 Hahn U, Bailey TM, Elvin LB. Effects of category diversity on learning, memory, and generalization. *Mem Cognit* 2005;33(2):289–302
- 25 Balslev T, Jarodzka H, Holmqvist K, et al. Visual expertise in paediatric neurology. *Eur J Paediatr Neurol* 2012;16(2):161–166
- 26 Klein GA, Hoffman RR. Seeing the invisible: Perceptual-cognitive aspects of expertise. In: Rabinowitz M, ed. *Cognitive Science Foundations of Instruction*. Hillsdale, NJ: Erlbaum; 1993:203–226
- 27 Lowe DG. *Perceptual Organization and Visual Recognition*. Boston, MA: Kluwer; 1985
- 28 Kennedy RA, Anderson-Montoya BL, Scerbo MW, et al. The influence of visual aids on detecting early and late decelerations in maternal-fetal heart rate patterns. In: *Proceedings of the 56th Annual Meeting of Human Factors & Ergonomics Society*; October 22–26, 2012; Boston, MA; 2012:1317–1321
- 29 Wertheimer M. Laws of organization in perceptual forms. In: Ellis WD, ed. *A Source Book of Gestalt Psychology*. London, United Kingdom: Kegan Paul, Trench, Trubner & Company; 1938:71–88