SEEING WEATHER THROUGH CHAOS

A Case Study of Disembedding Skills in Undergraduate Meteorology Students

PEGGY McNEAL, HEATHER PETCOVIC, TERESA BALS-ELSHOLZ, AND TODD ELLIS

Recognizing how students approach pattern recognition and where they struggle is critical to effective instruction, as how capably they discern patterns in complex data displays may enhance or hinder success in meteorology courses.

Meteorology coursework involves learning to analyze the spatial patterns of variables on surface and upper-air maps, such as temperature and pressure gradients. Acquiring skills to identify subtle signals and patterns through hand analysis remains foundational for learning to forecast, even

with modern numerical weather prediction. However, differences in proficiency, experience, and cognition contribute to the subjective nature of forecasting. For example, the ability to isolate and attend to one aspect of a complex display, a spatial ability particularly relevant to the interpretation of meteorological maps, is variable among individuals. Cognitive scientists and spatial thinking researchers refer to this ability as flexibility of closure or disembedding. It is described as “the ability to hold a given visual percept or configuration in mind, so as to disembed it from other well-defined perceptual material” (Ekstrom et al. 1976, p. 9). Disembedding ability can be measured with the Educational Testing Service’s (ETS) Hidden Figures Test (Fig. 1).

Multiple spatial thinking skills are described in the literature; however, we identified disembedding as a particularly important spatial skill used by meteorologists in our previous work (McNeal et al. 2018, 2019). Using a survey, we asked meteorologists to select, among six spatial thinking skills, those they perceive they use. The results pointed toward disembedding; 72% of the meteorologists in the sample ($N = 93$) reported using disembedding while interacting with nine meteorological charts and images (McNeal et al. 2018). We followed this with

AFFILIATIONS: McNEAL—Department of Physics, Astronomy and Geosciences, Towson University, Towson, Maryland; PETCOVIC—Mallinson Institute for Science Education, and Department of Geological and Environmental Sciences, Western Michigan University, Kalamazoo, Michigan; BALS-ELSHOLZ—Department of Geography and Meteorology, Valparaiso University, Valparaiso, Indiana; ELLIS—Mallinson Institute for Science Education, and Department of Geography, Western Michigan University, Kalamazoo, Michigan
CORRESPONDING AUTHOR: Peggy McNeal, pmcneal@towson.edu

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/BAMS-D-18-0015.1

A supplement to this article is available online (10.1175/BAMS-D-18-0015.2)

In final form 16 November 2018

©2019 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

a separate study of 81 novice through expert meteorologists. We measured participants' disembedding and mental rotation abilities, visuospatial working memory capacity, meteorology knowledge, and expertise and compared these scores to performance on a series of novice-level tasks similar to problems typically assigned in undergraduate meteorology courses (see online supplement at <https://doi.org/10.1175/BAMS-D-18-0015.2>). A regression analysis demonstrated that meteorology knowledge [$\Delta R^2 = 0.284$, $\Delta F(1, 76) = 30.21$, $p < 0.001$] and disembedding ability [$\Delta R^2 = 0.054$, $\Delta F(1, 75) = 19.16$, $p < 0.001$] significantly predicted performance on the meteorological tasks. Additionally, we regressed task performance onto disembedding ability categorized across knowledge levels to reveal an interesting finding: disembedding ability continued to have a significant effect on performance at both low and high levels of meteorology knowledge [$(M_{\text{low}} = 21.26$, $SD = 4.20$; $M_{\text{high}} = 30.79$, $SD = 2.04$), $t(78) = -12.76$, $p < 0.001$]. In other words, disembedding ability continued to affect the performance of even the expert meteorologists in our sample (McNeal et al. 2019).

Although this previous work identified disembedding as significant to performance on meteorology tasks, it did not yield insight into how meteorology students use disembedding skills and multiple questions remained. We were curious to learn what the process of disembedding looks like in a meteorological context and how it varies between high- and low-ability disembedders. We sought to identify patterns that are particularly difficult for students to disembed. We were also interested in revealing strategies that low-ability disembedders use when struggling to see patterns in meteorological data, and if these alternate strategies are effective. Making these processes visible, we reasoned, would be beneficial to

the meteorology community, especially meteorology instructors.

The importance of spatial skills, such as disembedding, in science, technology, engineering, and math (STEM) domains is well established (Uttal and Cohen 2012). A 20-yr, longitudinal study conducted by Shea et al. (2001) collected baseline spatial assessment data from 563 mathematically precocious 13-yr-old students, then collected additional biographical, educational, and occupational data at 5-, 10-, and 20-yr intervals. They found that those students who ultimately ended up in STEM careers 20 years later had higher levels of spatial ability at age 13. They concluded that spatial ability is a crucial component of human cognition that cannot be neglected in education, an argument echoed by recent researchers (e.g., Uttal and Cohen 2012; Uttal et al. 2013; Hegarty 2014). Additional research demonstrating the trainability of spatial ability makes this conclusion even more compelling (Uttal et al. 2013).

Thus, the current investigation used a case study approach to delve deeply into how three students with varying disembedding ability solve meteorology tasks. The purpose of the study was to “make visible” how these students use disembedding versus alternate strategies when approaching typical meteorology tasks encountered in undergraduate coursework. We aimed to highlight thought processes and approaches in order to better understand how students use this spatial thinking skill. If described and documented, students experiencing difficulty with disembedding may be better recognized by instructors, who can offer targeted disembedding training or instruction in alternate strategies. Uttal and Cohen (2012) suggest that spatial thinking skills serve as an early filter in undergraduate STEM courses. Students with higher spatial ability are often able to circumvent

limits imposed by underdeveloped content knowledge whereas students with lower spatial abilities may experience difficulty and give up (Hambrick et al. 2012). Therefore, understanding how students use disembedding in their meteorology coursework is an important component of understanding what makes students successful, scaffolding student learning, and maintaining students in meteorology programs.

FIG. 1. Sample item from the Hidden Figures Test (Ekstrom et al. 1976). In this test, individuals search the perceptual fields (I and II) to find one of the given configurations (A through E).

FIG. 2. Surface map for 0000 UTC 2 Mar 2018. Map source: Plymouth State Weather Center, used with permission.

AN EXAMPLE ILLUSTRATING DISEMBEDDING. We can exemplify the use of disembedding through the partial analysis of a surface map by an imaginary meteorologist. Figure 2 depicts the 1–3 March 2018 nor'easter at 0000 UTC 2 March 2018 and was obtained from the Plymouth State Weather Center. In this seemingly chaotic map, our meteorologist is able to extract layers of information by disembedding multiple patterns. For example, individual stations represent percent cloud cover with a shaded circle. By focusing attention solely on the shaded circles, the meteorologist can identify large-scale cloud patterns in both the western and eastern regions of the United States. This identification further guides the meteorologist's disembedding of additional patterns. Station wind barbs indicate wind direction, so next the meteorologist concentrates exclusively on the wind barbs, disembedding counter-clockwise circulation

patterns around low pressure centers and clockwise circulation patterns around high pressure centers. This gives the meteorologist an understanding of wind flow and a sense of where low and high pressure areas exist. In the case of the low centered over southwestern Pennsylvania (Fig. 3), the meteorologist next identifies a cold front associated with the low by selectively attending to the wind barbs again in Fig. 2 and disembedding marked wind shift patterns on either side. This pattern is observable along the Atlantic coast south of Virginia, extending to the southwest across northern Florida, and is annotated

in Fig. 3 (map analysis added by T. Ellis). The meteorologist will continue by disembedding additional patterns using pressure, temperature, and wind speed data, along with disembedding patterns in accompanying upper-air maps.

How our meteorologist—or any individual—processes visual information in order to perceive

FIG. 3. Surface analysis for 0000 UTC 2 Mar 2018 with pressure and fronts annotated. Map source: Plymouth State Weather Center, used with permission.

patterns is a matter of selectively keeping an eye on specified features while effectively ignoring all others. For example, when looking for cloud patterns, the meteorologist knows that the darkest circles indicate areas of complete cloud cover, and naturally focuses on darker regions without the need to conscientiously ignore the chaotic background. Once large regions are identified, our meteorologist maintains this focus to compare regions, seek boundaries, and visualize overall configurations.

Presumably, maintaining a focus on wind barbs, which have less visual salience than cloud cover data, is more difficult. Seeing this synoptically likely involves conceptually identifying what to look for, establishing a targeted focus, and mentally blocking everything else out. With this “lens,” our meteorologist may no longer see the chaos inherent in the maps, but mentally hunt for the targeted feature (in this case wind barbs) while simultaneously stitching together an overall picture. The meteorologist may also visually “zoom in and out” to bring the picture to the forefront of the mind, while continuing to “blur out” the competing noise. How our meteorologist does this—purposefully reveal sought-after meteorological patterns while effectively rendering the remaining chaos invisible—illustrates the essence of disembedding in a meteorological context.

CAPTURING STUDENT DISEMBEDDING.

In conjunction with the larger study¹ described above, we administered to $N = 81$ participants the Hidden Figures Test to measure level of disembedding ability; the Fundamentals in Meteorology Inventory, version 1.5 (FMI), to measure level of meteorology knowledge (Davenport et al. 2015); a domain experience questionnaire (DEQ) modified from two prior studies to measure level of meteorology experience (Baker et al. 2012; Petcovic et al. 2016); and a series of meteorology tasks, based on multiple versions of case study problems typically provided to undergraduate meteorology students. The meteorology tasks contained maps available on the Plymouth State Weather Center website (Plymouth State University 2018) and included a surface map with instructions to mark a low pressure center and draw warm and cold fronts. Upper-air maps representing 850, 500, and 300 hPa included instructions to annotate troughs and ridges, shade areas of cold and warm air advection (CAA/WAA), positive (cyclonic) and negative (anticyclonic)

¹ The study was approved by the Institutional Review Boards at Western Michigan University and Valparaiso University for research with human subjects. All participants signed informed consent documents.

THE THINK-ALoud METHOD FOR VERBAL DATA COLLECTION

The think-aloud method (e.g., Someren et al. 1994) was developed from older introspection methods established in conjunction with psychological research in the early 1900s. Duncker and Lees (1945) and De Groot (1946, 1965) were among early researchers who successfully used the think-aloud method in their work characterizing expertise in chess. Newell and Simon (1972) promoted the method by showing that detailed verbal explanations and data are reliably obtainable. The method is today commonly used in psychology research, problem-solving research, and in the development of artificial intelligence (Someren et al. 1994). In addition to its use characterizing expert–novice problem-solving in physics by Chi (1981), the method has been used successfully in numerous geoscience education research studies (e.g., Callahan 2013; Dickerson et al. 2005; Jolley et al. 2013; Petcovic et al. 2009) as well as investigations of weather forecasting ability (e.g.,

Joslyn and Jones 2006; Trafton et al. 2000).

The think-aloud method requires subjects to continuously speak aloud whatever thoughts come to mind while performing a task. There are no interruptions, suggestive prompts, or questions, greatly enhancing the validity of the data in capturing moment-by-moment thinking. We collected think-aloud data as participants completed the series of meteorological tasks. Before beginning the think-aloud, we warmed up participants using a script with instructions and a practice scenario obtained from a study (Tenbrink and Taylor 2015) that followed methods outlined by Ericsson and Simon (1998). This opportunity to practice with an unrelated scenario allowed the participants to become familiar with the think-aloud process while giving us an opportunity to train the participant to continuously verbalize their thoughts. Someren

et al. (1994) note that the think-aloud method has some limitations since it disrupts mental processing even as it attempts to capture it authentically. Affective factors, such as embarrassment over a knowledge deficit, can result in cognitive processes different from those when thoughts remain internal. Moreover, talking takes longer than thinking and synchronization can be an issue leading to silent periods that lack interpretation. For complex tasks that already tax working memory capacity, thinking aloud can add to cognitive load. However, meteorologists (and meteorologists in training) become accustomed to giving forecast briefings, which from the standpoint of psychology are a natural analog to the think-aloud method (Hoffman et al. 2017). We propose that this diminishes some concern with using the think-aloud method in our study and reiterate that the method has a robust history in established research.

vorticity advection (PVA/NVA), and divergence and convergence. Finally, given a U.S. base map, we asked participants to mark the location of lowest pressure 12 h later. It took participants 15–30 min to complete the task series. We scored each completed map with a 10-point rubric to produce a meteorology task score.

From this sample of practicing meteorologists, meteorology professors, and student meteorologists, we solicited volunteers who participated in a think-aloud (see sidebar “The think-aloud method for verbal data collection”) while they completed the meteorology tasks. This produced a sample of 10 participants [one U.S. Navy aerographer’s mate, one National Weather Service (NWS) forecaster, two professors, and six meteorology students] who provided verbal data. We video recorded the think-aloud process using a camera and tripod that captured the participant’s voice, along with hand gestures and annotations on the weather maps. We transcribed the audio portion of the recording and used the video and participant products for interpretation of the transcribed data.

In the current study, we focused on the following research question: How do three meteorology students of variable disembedding ability (low, medium, and high) approach and complete a series of meteorology tasks? To address this question, we used a qualitative case study design and pursued in-depth analyses of three cases (see sidebar “Case studies in qualitative social science research”). We purposefully selected data (think-aloud videos, transcripts, and participant products) from three meteorology students who were similar in age and expertise yet had high, medium, and low disembedding skill per their scores on the Hidden Figures Test. All were undergraduates at the same midwestern university, within a large meteorology program. Each participant was taking upper-level meteorology coursework.

We analyzed the data using both emergent and a priori coding schemes (both coding schemes are included in the online supplement). We developed the a priori codes by considering what patterns a meteorologist is expected to “see” within the context of each map, (i.e., the data that a meteorologist needs to disembed in order to successfully complete the task at hand). Thus, the a priori codes were intended to capture instances of participant disembedding.

The second set of codes emerged through open-ended initial coding (Saldaña 2015), and uncovered what participants were doing when they were not disembedding. Of these, two codes described indiscriminate identification of features. For example, by recognizing areas where “boxes” are created

CASE STUDIES IN QUALITATIVE SOCIAL SCIENCE RESEARCH

Similar to case studies of meteorological events, qualitative social science research uses case study methods to investigate single cases comprehensively. Purposefully small sample sizes allow researchers to collect exquisite detail about individual cases and conduct cross-case theme analysis (Creswell 2013). Like the careful analysis of a single meteorological event, the defining feature of a qualitative case study is the unit of analysis (Merriam and Tisdell 2015). Thus, the power of qualitative case study research lies in its ability to deeply probe the unique behaviors and experiences of an individual and compare to other distinct cases. Comparable to in-depth understanding of a single storm, this increases knowledge of human behavior (such as how disembedding manifests during problem-solving) without generalizing results to the larger population. Case study research has a long and distinguished history across many disciplines (Creswell 2013), notably in education research.

with the 850-hPa height contours and isotherms, a participant can pinpoint areas of cold and warm air advection. In comparison, participants who lacked this pattern-recognition skill instead identified broad areas east and west of a trough or broad areas along warm and cold fronts. A third emergent code captured explicit expressions indicating difficulty resolving patterns, such as “it’s kind of muddled up there.”

Qualitative research must meet rigorous standards of trustworthiness that establish the equivalence of reliability in quantitative studies. Researchers rely on several methods, alone or in combination, to ensure trustworthiness [see Merriam and Tisdell (2015) for a full discussion]. For the interested reader, we include our methods for establishing trustworthiness in the online supplement.

OBSERVING HIGH, MEDIUM, AND LOW DISEMBEDDERS. We gave our three undergraduate meteorology participants (all males) the pseudonyms Abe, Bart, and Cole (Table 1). They represent, in order, low, medium, and high disembedding skill. Our focus in the analysis was on *how* these participants approached and completed the task through disembedding, regardless of their ultimate success at the task.

Abe (low disembedder) is 20 years old and has completed courses in weather analysis, dynamic meteorology, and forecasting. He enjoys storm chasing and has strong interests in severe and winter weather as well as in computer models, programming, and remote sensing.

TABLE 1. Participant information, $n = 3$. The data demonstrate that Bart and Cole have much closer Hidden Figures Test scores than either have to Abe, and this is evident in the results. All participants' DEQ scores fall in the bottom 20% of scores represented in our larger study.

Name	Hidden Figures Test score	Disembedding level	DEQ score	Meteorology task score	FMI score	Time spent completing meteorology tasks
Abe	4	Low	0.75	21	19	33.0 min
Bart	10	Medium	1.00	19	24	20.5 min
Cole	12	High	2.25	20	31	22.0 min
Range ($N = 81$)	1–16		0.75–11.75	9–35	11–34	

Bart (medium disembedder) is also 20 years old and has completed courses in weather analysis, synoptic meteorology, and physical meteorology. Bart's experience growing up in the Midwest exposed him to several weather extremes, including tornados, blizzards, lake effect snow, and derechos.

Cole (high disembedder) is 21 years old and has completed courses in weather analysis, synoptic and dynamic meteorology, physical meteorology, and forecasting. Like Bart, his experience growing up in the Midwest, especially experiences with severe weather, sparked his interest in weather. His higher DEQ score is the result of an internship.

We present results by map type, working from the surface map and moving through the upper-air maps representing conditions at 850, 500, and 300 hPa. For each, we begin with Cole, who had the highest level of

disembedding skill, followed by Bart (medium level), then Abe (low level).

The surface map. To find the center of low pressure, Cole first sought counter-clockwise rotation and noted, "It's probably in here somewhere." He inspected station pressure readings next and remarked, "So it looks like that's the low right here." He marked a center of low pressure but repositioned it later after discovering a lower pressure, saying, "So that would be, the lowest pressure now." To locate warm and cold fronts, Cole looked for changes in wind direction, temperature, and dewpoint: "Change in wind direction is here and temperature and dew point. So, it would probably be down here in between that 75 and 65 where the dew point begins to drop." Cole spent 6 min with this process,

FIG. 4. Surface analyses with low pressure and fronts annotated.

carefully checking the data and going back over it (Fig. 4).

Bart extracted a mix of data from multiple station plots to place his low and fronts simultaneously. He considered changes in wind direction before isolating temperature and pressure data: “You see winds feeding into it in a counter-clockwise motion. You have 1,000.7 mb right here at this station plot, which appears to be just about the lowest that I can see.” He found warm and cold fronts by locating steep temperature gradients with temperature data from multiple station plots: “So, looking at temperature plots, we have 48 and 59 within a small gradient.” Like Cole, he spent 6 min on this task (Fig. 4).

Abe struggled with the surface map and spent 10.5 min grappling with it. He made multiple references to difficulty “seeing” data and patterns in the map:

- “Oh boy, it gets really cluttered up here.”
- “I’m not seeing a really large shift in winds here.”
- “I’m just going to try to identify the area of this low. According to the surface map, it’s kinda hard to tell where that would be.”
- “I’m having a hard time kind of seeing that.”
- “It’s kind of muddled up there. This is a tough map.”

Abe only used wind direction to place fronts and the low. He drew fronts first, then referred to the 850- and 500-hPa analyses (completed prior) for

help in locating the surface low, but these did not correspond to his frontal placement on the surface analysis. Abe placed the low in Missouri, “because I do have this sort of circulatory area,” however, this offset the low from the crest of the fronts. He knew this was incorrect and seemed uncomfortable, but in the end, resigned and stated, “I’m not seeing anything else that’s popping out.” (Fig. 4)

The 850-hPa map. After marking a trough and two ridges on the 850-hPa map, Cole located an area of CAA, stating, “Cold air advection you would have down here because you’re crossing your wind heights across the temperature lines.” He sought additional regions of CAA: “We still have it crossing...the wind blowing into the high values.” Next, Cole identified a small pocket of WAA and commented, “It looks like a little bit here. You sort of have the winds going across, blowing warmer temperatures in.” He initially limited his shading to a small region (Fig. 5), then noted, “Then temperature and height lines are parallel, so you don’t have any advection until we get up here where they start crossing again.”

From analysis of Bart’s think-aloud, we conclude that while marking a trough and ridge, Bart misinterpreted isotherms as height contours, hence the placement seen in Fig. 5. Bart used wind direction to identify areas of thermal advection. He described his identification of WAA:

FIG. 5. 850-hPa analyses with WAA shaded in red and CAA shaded in blue.

We're going to shade in just about everywhere that's southerly flow. So, start down here pretty much all through Florida and through the Gulf, up north through the Southeast. And then keep going. The wind kind of starts to get westerly as soon as you get to Illinois so we're going to start fading east more. And you still have that southerly flow pretty much all the way through to the top of this ridge here. So, pretty much along through the East Coast here and then all the way down through the Southeast and into the Gulf. That's where I have this warm air advection.

Bart's treatment of CAA was the same; he isolated wind flow patterns to infer CAA:

Cold air advection, if you look by this trough area, there's northerly flow, coming down from central Canada, so it's going to bring some cold air advection. Um, so you have kind of a westerly flow here coming off the West Coast, so it's not going to be too much of a cold air advection situation. So, you have that northerly flow also going through the plains here and then kind of rounding off right about here in western Texas. So, I guess we'll start down there in western Texas and the northern part of New Mexico, and just go straight north and go a little bit east as you start to get into the plains. And then just straight up through the plains and into Canada. So, there's your cold air advection.

Thus, Bart identified WAA and CAA indiscriminately, using wind direction as the sole determinant of where he shaded.

Abe identified one trough and two ridges through height line curvature: "It's really all about just defining these areas of whether it's really distinctly curved or whether it's not." To identify CAA, Abe inspected isotherm spacing and indicated where "temperature gradients are very, very tight." He generalized, "I would imagine that this would be a cold front here. I'm going to go ahead and put cold air advection along this area of this cold front which is associated with this large extratropical cyclone." Abe's identification of warm air advection was similar. He relied on isotherm spacing, but additionally observed wind direction data, saying "You've got a lot of really strong southerly flow," gesturing how air would move to the north. Finally, he stated, "I'm going to go ahead and . . . circle those areas where it's very, very strong." He gave no rationale for how he identified these "strong" areas (Fig. 5).

The 500-hPa map. Our intent was for participants to infer curvature vorticity through height line curvature and shear vorticity through relative wind speed, and to determine advection also using wind speed. However, these three participants only considered curvature vorticity. It was apparent that they were less confident with vorticity advection and all completed the task

Abe's 500 hPa analysis
(low disembedder)

Bart's 500 hPa analysis
(medium disembedder)

Cole's 500 hPa analysis
(high disembedder)

FIG. 6. 500-hPa analyses with PVA shaded in red and NVA shaded in blue.

quickly. Their completed analyses exhibit vorticity advection relative to troughs and ridges (Fig. 6); however, the rationale given by each participant gives us a glimpse into how they determined placement.

Cole identified one trough, then gestured with his finger east of the trough in a counter-clockwise motion while commenting, “You have vorticity at the trough; that would be positive vorticity advection.” He isolated wind speed data to determine advection: “You’d have your vorticity at the base, so it would advect with strong winds.” He spent less time contemplating NVA and placed it west of the trough: “So NVA down here, in that general area.”

Bart identified one trough and one ridge and based his designation of PVA and NVA entirely in relation to the trough. While shading red, he stated, “So, here you have your vorticity max at the base of this trough, so that means that your PVA is going to be right here, just east, northeast of that trough base.” Similarly, as he shaded blue, he said, “And your negative vorticity advection is going to be just west of that trough base as you approach that subgeostrophic flow down there and as you approach that vort max. So, there’s going to be your negative vorticity advection.”

Abe identified a trough and multiple ridges and based his designation of PVA and NVA in relation to these. He expressed less familiarity with the concept, applied rule-based reasoning, and indicated a preference for model interpretation:

- “I know that in the right exit region, over here, I would imagine that we would get a large area of positive vorticity advection and that’s just kind of a rule that I learned a few weeks ago in my MET 300 class.”
- “This is something that I definitely rely on the models for is identifying areas of vorticity.”
- “I think that’s about all that I would be comfortable with for that.”

The 300-hPa map. Our intent was for participants to identify diffluence and confluence through the spreading or converging of height contours and divergence and convergence through the horizontal speeding up or slowing down of winds. These three participants, however, exhibited alternate approaches.

Cole sought the presence of a jet streak, pointing out the closely spaced height contours with his finger:

So, you’d have divergence in the exit region of where your jet streak would be, which would be here. We have the stronger winds and the tight packing of the height lines. It looks like the strongest winds in through here, so you probably have your divergence in this area.

He referenced speed convergence: “Down here is where you’d have your convergence. That would

Abe's 300 hPa analysis
(low disembedder)

Bart's 300 hPa analysis
(medium disembedder)

Broadly shaded blue area
representing region of
convergence

Dashed line indicating trough

Broadly shaded red area
representing region of
divergence

Cole's 300 hPa analysis
(high disembedder)

FIG. 7. 300-hPa analyses with divergence shaded in red and convergence shaded in blue.

make sense when you're in the base of the trough. So, convergence here, which makes sense because the winds would be backing up." It is unclear whether Cole was using knowledge about the movement of air through a trough or disembedding wind speed data from the map; the former is likely (Fig. 7).

Bart made quick work of the 300-hPa map by applying knowledge of the behavior of air moving through troughs and ridges and broadly shading areas of divergence and convergence accordingly:

So, your upper-level divergence is going to be right here, right as you start to approach, as the wind starts to get more geostrophic and approach that supergeostrophic flow at the top of the ridge. So, our upper-level divergence is going to be right here through...the western Midwest and the northern part of Texas, right around there. Our upper-level convergence is going to occur right here in the west as you approach that subgeostrophic flow at the trough base. That wind's going to start to slow down, so we're going to have that upper-level convergence right here through California and Nevada and into the western Rocky Mountains. And there's your upper-level convergence.

Thus, we see no evidence of disembedding with Bart on the 300-hPa map (Fig. 7); rather, he applied his knowledge and rule-based reasoning.

Abe's 300-hPa map (Fig. 7) is interesting because he was especially conscientious about identifying multiple troughs and ridges in the jet stream; this exhibits disembedding skill. With this accomplished, however, Abe used very simplistic rules of thumb to identify divergence and convergence: "Convergence is usually found on the left side of these trough lines, so I'm going to go ahead and just shade that in right there. And then on the other side, we've got divergence."

CHARACTERIZING THREE STUDENT APPROACHES. To support and illustrate findings from the verbal data, we counted coded units of text and grouped them as 1) instances of disembedding, 2) alternatives to disembedding, and 3) expressions of difficulty. We expressed each as a percent of the total of each participant's coded verbal data and displayed the data graphically by participant (Fig. 8).

Figure 8 illustrates an increasing use of disembedding and a decreasing reliance on alternatives to disembedding with increased disembedding skill while completing the meteorology tasks.

Alternatives to disembedding include broad identification of features relative to troughs, ridges, and fronts, thus capturing common rules of thumb, a reasonable strategy for completing the tasks. All three participants used both disembedding and alternatives to disembedding; however, the relative amount employed by each is noteworthy. We propose that meteorology knowledge and rule-based reasoning are most effective when used in conjunction with disembedded weather features and patterns from data. This was evident in our results and suggests that Abe, with lower disembedding skills, may have been hindered when performing meteorology tasks such as the one we provided.

Percent of participants' total coded text by category

FIG. 8. Code assignment by percent of participants' total codes.

Analysis of the think-aloud data allowed us to answer our research question and characterize how three meteorology students of variable disembedding ability approached and completed a series of meteorology tasks. Beginning with Abe (low disembedder), the only data he attended to on the surface map was wind direction. This, coupled with eight comments indicating difficulty discerning station plot data, demonstrates that he struggled to identify relevant information from the distracting background. With the 850-hPa map he used both temperature and wind information to identify areas of thermal advection and isolated areas with tight temperature gradients. He did not take the next step, however, and seek the intersection of height contours and isotherms. Abe was keen on identifying every bulge and kink in the contours on the 500- and 300-hPa maps, which suggests a level of disembedding skill, at least with this feature. His straightforward shading of PVA/NVA and divergence/convergence in relation to troughs and ridges was essentially correct, but relied solely on rules of thumb rather than pattern recognition. We characterize his task processing as a straightforward interpretation of salient features with a reliance on rules of thumb as an accommodation for limitations with disembedding.

Bart (medium disembedder) also broadly shaded PVA/NVA and divergence/convergence in relation to troughs and ridges. He augmented identification of divergence and convergence with recognition of speed convergence; however, this may have stemmed from knowledge of jet stream airflow rather than wind information embedded in the map. This approach is a practical accommodation that does not rely on disembedding. Bart disembedded wind patterns on the 850-hPa map, but without thermal patterns, his shading of WAA and CAA was generalized and included areas unaffected by CAA and WAA. His identification of surface features stemmed from observations of temperatures and pressures from multiple stations, along with tendencies for each. We characterize Bart's disembedding as further progressed than Abe's, due to his increased interrogation of the maps, particularly the surface map, which resulted in recognizing patterns associated with the warm and cold fronts and a more accurate placement of each.

Cole's (high disembedder) analysis was characterized by instances of attending to embedded aspects in the maps that were not identified by the other participants. With the 850-hPa map, he disembedded the intersection of isotherms and height contours to discern precise regions of cold and warm air advection. On the 500-hPa map he referenced curvature

and identified areas of strong winds that would support vorticity advection. On the 300-hPa map he resolved wind patterns within the jet stream, attempting to identify the existence of a jet streak. He then engaged his knowledge of jet streak dynamics, an example of rule-based reasoning. In this case, we note that Cole used higher-level disembedding and rule-based reasoning together, and we recognize this as a further sophistication of skill. Overall, Cole demonstrated purposeful pattern seeking, an enhanced ability to extricate patterns from the complex graphics displayed in the maps, and a capacity for combining rule-based reasoning with information gleaned from patterns.

These results exemplify the intertwined nature of meteorology knowledge and disembedding as part of complex data interpretation. For meteorology instructors, it is important to recognize this relationship and identify when students struggle with either. Lack of meteorology knowledge is apparent; identifying students who struggle with disembedding is more challenging. Illustrating common strengths and weaknesses, as we have done here, will hopefully begin to fill this gap, while further investigation is necessary to continue this line of research.

CONCLUSIONS. Individuals rely on different repertoires of experience, knowledge, and cognitive skills to solve problems. We cannot entirely discern to what degree the differences we identified were due to disembedding ability or prompted by meteorology knowledge. Presumably, you cannot disembed patterns that you do not know to look for. Additionally, our quantitative analysis indicates that knowledge is the primary predictor of success with the meteorology tasks. However, the results of our qualitative case study analyses indicate that disembedding skill did factor into how each student approached and completed the meteorology tasks and that disembedding skill merits consideration in undergraduate meteorology training.

LIMITATIONS AND IMPLICATIONS FOR METEOROLOGY EDUCATION. Our results provide an in-depth characterization of three individuals in a single university setting. Thus, we refrain from generalizing findings to other students with similar disembedding capabilities, who possibly develop different approaches. However, we believe that one of the most productive applications of this research is to promote instructor awareness of students' disembedding skills in undergraduate meteorology classrooms. Recognizing and addressing

students who struggle while extracting information and looking for patterns in meteorological data has potentially been a neglected aspect of student formative assessment—that is, assessment intended to generate feedback on performance to improve and accelerate learning (Sadler 1998). We imagine that such struggling students may hesitate to make their difficulties known, especially in a classroom full of students for whom the skill seems natural (or appears to). Additionally, meteorology instructors, for whom such tasks come easy, may be unaware of the challenges faced by students with low disembedding ability and attribute difficulty to other factors. Explicitly recognizing the student disembedding process—what it looks like, where students have trouble, and what types of accommodations they use—can inform targeted instruction and scaffold student learning.

Such instruction hinges on the idea that we can train students to be better spatial thinkers, a premise grounded in robust research-based evidence. A meta-analysis of 217 spatial training studies completed by Uttal et al. (2013) demonstrates that spatial training improves spatial thinking skills with an average effect size (Hedges's *g*) of 0.47 when compared to control groups. Analyses of studies with delayed post-testing indicate that spatial training effects endure through post-test time periods of up to a month. Training in spatial thinking also appears to transfer to spatial tasks other than those used in training—both to similar tasks and tasks that require different types of spatial skills or representations (Uttal et al. 2013).

Uttal et al. (2013) calculated an effect size for a subsample of studies that trained intrinsic-static spatial skills (see Newcombe and Shipley 2015), of which disembedding is one, and calculated an effect size (Hedges's *g*) of 0.32, indicating a lower training effect with intrinsic-static skills. Not only are theories addressing this difference absent, but studies specifically investigating disembedding—its nature, extent, and potential training—are lacking in the spatial thinking literature, which focuses principally on spatial transformations, such as mental rotation. Further investigation is needed to better understand the trainability of intrinsic-static spatial skills such as disembedding and how to teach these skills most effectively.

Of those papers that address disembedding in STEM (science, technology, engineering, and mathematics) learning, Reynolds (2012) discusses disembedding as a key competency for geology students who interpret complex outcrops, patterns on geologic maps, aerial photos, and satellite images.

He shares his personal experience teaching geology, in which he includes explicit disembedding instruction. He accomplishes this by highlighting observable examples, then prompts students to describe and distinguish visible features before modeling how to mentally and visually isolate one from another. This approach could potentially be used in meteorology classes, highlighting, for example, how to disembed data from a surface chart in order to locate fronts.

Atit et al. (2016) investigated novice interpretation of contour lines on topographic maps, a representation we consider analogous to isopleths and disembedding patterns on surface and upper-air maps. They found that the use of pointing and tracing gestures with contours helped novice map users to associate contour lines with the elevation information they encode. Additionally, specific vocabulary (e.g., shallow, steep, slope), when coupled with pointing and tracing gestures, facilitated interpreting the meaning of the contour lines in terms of elevation and thinking about the shape of the represented terrain.

Expanding beyond disembedding, two additional examples illustrate how spatial training benefits students in other STEM education domains. Sorby and Baartmans (2000) recognized that spatial visualization of engineering problems, specifically the ability to mentally rotate objects, is essential to student success in foundational engineering courses. Thus, they developed a pre-graphics course for freshman engineering majors who demonstrate weak mental rotation ability on standard spatial thinking tests. The course has had a positive impact on developing student spatial skills, improving student grades in follow-on engineering courses, and improving retention rates (Sorby 2007). Ormand et al. (2017) developed curricular materials for structural geology courses that develop penetrative thinking ability necessary to visualizing folded and overturned rock units. The use of exercises to promote penetrative thinking has shown significant increases in student ability to solve geological problems requiring such skills (Ormand et al. 2017). These successes exemplify the potential for enhancement of current meteorology education and training.

Baenninger and Newcombe (1989) outline two general types of spatial training. The simplest is task specific and provides repeated exposure to items on spatial thinking tests followed by assessing transfer through improvement on other spatial ability measures. A second type of spatial training offers instruction beyond practice with test items. A range of possibilities includes coaching test item responses, demonstrating and modeling applied

spatial thought (e.g., Reynolds 2012), and providing prolonged exposure to activities or coursework shown to improve spatial ability (e.g., Sorby and Baartmans 2000; Ormand et al. 2017). We propose that similar opportunities are ripe for further development and evaluation in undergraduate meteorology classrooms and encourage a creative dialog about collaborative efforts moving forward.

ACKNOWLEDGMENTS. We thank the Graduate College at Western Michigan University for funding our research and Dr. David Changnon and Dr. Walker Ashley for facilitating access to participants. Special thanks go to Dr. Nicole LaDue and three anonymous reviewers for providing feedback on the manuscript, and to our participants for agreeing to let us observe and record them.

REFERENCES

- Atit, K., S. M. Weisberg, N. S. Newcombe, and T. F. Shipley, 2016: Learning to interpret topographic maps: Understanding layered spatial information. *Cognit. Res. Princ. Implic.*, **1**, 2, <https://doi.org/10.1186/s41235-016-0002-y>.
- Baenninger, M., and N. Newcombe, 1989: The role of experience in spatial test performance: A meta-analysis. *Sex Roles*, **20**, 327–344, <https://doi.org/10.1007/BF00287729>.
- Baker, K. M., H. L. Petcovic, M. Wisniewska, and J. C. Libarkin, 2012: Spatial signatures of mapping expertise among field geologists. *Cartogr. Geogr. Inf. Sci.*, **39**, 119–132, <https://doi.org/10.1559/15230406393119>.
- Callahan, C. N., 2013: An embodied perspective on expertise in solving the problem of making a geologic map. Ph.D. dissertation, Western Michigan University, 555 pp., <https://scholarworks.wmich.edu/dissertations/186>.
- Campbell, J. L., C. Quincy, J. Osserman, and O. K. Pedersen, 2013: Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociol. Methods Res.*, **42**, 294–320, <https://doi.org/10.1177/0049124113500475>.
- Chi, M. T., P. J. Feltovich, and R. Glaser, 1981: Categorization and representation of physics problems by experts and novices. *Cognit. Sci.*, **5**, 121–152, https://doi.org/10.1207/s15516709cog0502_2.
- Creswell, J. W., 2013: *Qualitative Inquiry and Research Design*. 3rd ed. Sage Publications, 472 pp.
- Davenport, C. E., C. S. Wohlwend, and T. L. Koehler, 2015: Motivation for and development of a standardized introductory meteorology assessment exam. *Bull. Amer. Meteor. Soc.*, **96**, 305–312, <https://doi.org/10.1175/BAMS-D-13-00157.1>.
- De Groot, A. D., 1946: Het denken van den schaker: Een experimenteel-psychologische studie. Noord-Hollandsche Uitgevers Maatschappij, 315 pp.
- , 1965: *Thought and Choice in Chess*. Walter de Gruyter GmbH, 482 pp.
- Dickerson, D., T. J. Callahan, M. Van Sickle, and G. Hay, 2005: Students' conceptions of scale regarding groundwater. *J. Geosci. Educ.*, **53**, 374–380, <https://doi.org/10.5408/1089-9995-53.4.374>.
- Dunker, K., and L. S. Lees, 1945: On problem-solving. *Psychol. Monogr.*, **58**, i–113, <https://doi.org/10.1037/h0093599>.
- Ekstrom, R. B., J. W. French, H. H. Harman, and D. Derman, 1976: *Manual of the Kit of Factor-Referenced Cognitive Tests*. Educational Testing Service, 216 pp.
- Ericsson, K. A., and H. A. Simon, 1998: How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind Cult. Act.*, **5**, 178–186, https://doi.org/10.1207/s15327884mca0503_3.
- Hambrick, D. Z., and Coauthors, 2012: A test of the circumvention-of-limits hypothesis in scientific problem solving: The case of geological bedrock mapping. *J. Exp. Psychol. Gen.*, **141**, 397, <https://doi.org/10.1037/a0025927>.
- Hegarty, M., 2014: Spatial thinking in undergraduate science education. *Spat. Cognit. Comput.*, **14**, 142–167, <https://doi.org/10.1080/13875868.2014.889696>.
- Hoffman, R. R., J. G. Trafton, D. S. LaDue, P. J. Roebber, and H. M. Mogil, 2017: *Minding the Weather: How Expert Forecasters Think*. MIT Press, 488 pp.
- Jolley, A., F. Jones, and S. Harris, 2013: Measuring student knowledge of landscapes and their formation timespans. *J. Geosci. Educ.*, **61**, 240–251, <https://doi.org/10.5408/12-307.1>.
- Joslyn, S., and D. Jones, 2008: Strategies in naturalistic decision-making: A cognitive task analysis of naval weather forecasting. *Naturalistic Decision Making and Macrocognition*, J. M. Schraagen et al., Eds., Ashgate Publishing, 183–202.
- McNeal, P., T. Ellis, and H. Petcovic, 2018: Investigating the foundations of spatial thinking in meteorology. *J. Geosci. Educ.*, **66**, 246–257, <https://doi.org/10.1080/10899995.2018.1483119>.
- , H. Petcovic, N. LaDue, and T. Ellis, 2019: Identifying significant cognitive factors for practicing and learning meteorology. *J. Oper. Meteor.*, **7**, 1–26, <https://doi.org/10.15191/nwajom.2019.0701>.
- Merriam, S. B., and E. J. Tisdell, 2015: *Qualitative Research: A Guide to Design and Implementation*. 4th ed. John Wiley and Sons, 638 pp.
- Newcombe, N. S., and T. F. Shipley, 2015: Thinking about spatial thinking: New typology, new

- assessments. *Studying Visual and Spatial Reasoning for Design Creativity*, Springer, 179–192, https://doi.org/10.1007/978-94-017-9297-4_10.
- Newell, A., and H. A. Simon, 1972: *Human Problem Solving*. Prentice-Hall, 920 pp.
- Ormand, C. J., and Coauthors, 2017: The Spatial Thinking Workbook: A research-validated spatial skills curriculum for geology majors. *J. Geosci. Educ.*, **65**, 423–434, <https://doi.org/10.5408/16-210.1>.
- Petcovic, H. L., J. C. Libarkin, and K. M. Baker, 2009: An empirical methodology for investigating geocognition in the field. *J. Geosci. Educ.*, **57**, 316–328, <https://doi.org/10.5408/1.3544284>.
- , C. J. Ormand, and R. W. Krantz, 2016: Earth, mind, and paper: Field sketches as expert representations of the Hat Creek fault zone. *3-D Structural Interpretation: Earth, Mind, and Machine*, R. W. Krantz, C. J. Ormand, and B. Freeman, Eds., AAPG Memoir, No. 111, AAPG, 173–189.
- Plymouth State University, 2018: Make your own: Product generator for real-time and archived data. Plymouth State Weather Center, <http://vortex.plymouth.edu/myo/>.
- Reynolds, S. J., 2012: Some important aspects of spatial cognition in field geology. *Earth and Mind II: A Synthesis of Research on Thinking and Learning in the Geosciences*, K. A. Kastens and C. A. Manduca, Eds., GSA Special Paper, No. 486, Geological Society of America, 75–77, [https://doi.org/10.1130/2012.2486\(12\)](https://doi.org/10.1130/2012.2486(12)).
- Sadler, D. R., 1998: Formative assessment: Revisiting the territory. *Assess. Educ.*, **5**, 77–84, <https://doi.org/10.1080/0969595980050104>.
- Saldaña, J., 2015: *The Coding Manual for Qualitative Researchers*. 3rd ed. Sage Publications, 368 pp.
- Shea, D. L., D. Lubinski, and C. P. Benbow, 2001: Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *J. Educ. Psychol.*, **93**, 604–614, <https://doi.org/10.1037/0022-0663.93.3.604>.
- Someren, M. V., Y. F. Barnard, and J. A. Sandberg, 1994: *The Think Aloud Method: A Practical Approach to Modeling Cognitive Processes*. Academic Press, 208 pp.
- Sorby, S. A., 2007: Developing 3D spatial skills for engineering students. *Australas. J. Eng. Educ.*, **13**, 1–11, <https://doi.org/10.1080/22054952.2007.11463998>.
- Sorby, S. A., and B. J. Baartmans, 2000: The development and assessment of a course for enhancing the 3-D spatial visualization skills of first year engineering students. *J. Eng. Educ.*, **89**, 301–307, <https://doi.org/10.1002/j.2168-9830.2000.tb00529.x>.
- Tenbrink, T., and H. A. Taylor, 2015: Conceptual transformation and cognitive processes in Origami paper folding. *J. Probl. Solving*, **8**, 2–22, <https://doi.org/10.7771/1932-6246.1154>.
- Trafton, G. J., S. S. Kirschenbaum, T. L. Tsui, R. T. Miyamoto, J. A. Ballas, and P. D. Raymond, 2000: Turning pictures into numbers: Extracting and generating information from complex visualizations. *Int. J. Hum. Comput. Stud.*, **53**, 827–850, <https://doi.org/10.1006/ijhc.2000.0419>.
- Uttal, D. H., and C. A. Cohen, 2012: Spatial thinking and STEM education: When, why, and how? *Psychology of Learning and Motivation*, Vol. 57, B. H. Ross, Ed., Academic Press, 147–151, <https://doi.org/10.1016/B978-0-12-394293-7.00004-2>.
- , N. G. Meadow, E. Tipton, L. L. Hand, A. R. Alden, C. Warren, and N. S. Newcombe, 2013: The malleability of spatial skills: A meta-analysis of training studies. *Psychol. Bull.*, **139**, 352–402, <https://doi.org/10.1037/a0028446>.