

**Expanding the Measurement, Use, and Support of Spatial Reasoning in STEM and  
the Geosciences**

By

Elijah Thomas Johnson

A dissertation submitted to the Graduate Faculty of  
Auburn University  
in fulfillment of the  
requirements for the Degree of  
Doctor of Philosophy

Auburn, Alabama  
August 5, 2023

Keywords: Spatial thinking, geologic block models, structural geology, geoscience education  
research

Approved by

Karen S. McNeal, Chair, Professor of Geosciences, Auburn University  
Carol Ormand, Science Education and Research Associate, Science Education Resource Center,  
Carleton College  
Jordan Harshman, Associate Professor of Chemistry and Biochemistry, Auburn University  
Stephanie Shepherd, Associate Professor of Geosciences, Auburn University

## **Abstract**

Spatial thinking has been well-described and characterized for many STEM fields, and there have been numerous studies that emphasize the importance of incorporating spatial opportunities early and often throughout curricula. However, there is still a gap in the research in implementing deliberate spatial training in curricula as well as the transfer of spatial thinking skills from academic to workforce settings. The goals of this dissertation are three-fold. Firstly, develop a survey to understand adults' spatial reasoning, experience with spatial tasks, and confidence in spatial reasoning to determine how these metrics along with college major and career fields are related. Secondly, employ a spatially-rich geological activity to determine how the delivery method of pedagogical aides (e.g., physically vs virtually) affect performance on a topographical and geological activity and a geoscience spatial thinking instrument. Lastly, describe and characterize the challenges to and strategies for solving the spatial visualization items on this spatial instrument using eye-tracking data and retrospective interviews. This research highlights the value of instruction and collaboration in STEM and the geosciences.

## Acknowledgements

If I were to name everyone who has helped me in the many ways that I have needed throughout graduate school, these acknowledgments would be longer than the dissertation itself. However, I will attempt to highlight as many as I can in this limited space and hope that the rest of you will accept my sincerest appreciation for your support and grace.

Firstly, I would like to thank my committee for their dedication and commitment to bringing out the best in my work and myself as a researcher, educator, and student. A special thanks is for my advisor and committee chair, Dr. Karen McNeal, for her coordination, expertise, and leadership throughout my graduate program in both my Master's and PhD. To the rest of my committee, Drs. Carol Ormand, Jordan Harshman, and Stephanie Shepherd, as well as my external reader, Dr. Cissy Ballen, thank you for your patience and enthusiasm. You all have made me feel like I deserve to be here and that all of my achievements matter. I would also like to thank Dr. Kim Mulligan for being my confidant, mentor, and cheerleader for all things academic and beyond.

I would also like to thank the faculty and staff of the Department of Geosciences at Auburn University and the Department of Chemistry at Auburn University at Montgomery for the opportunities you have given me to teach, support research projects, and volunteer. Thank you to my students who I have had the pleasure to instruct during these last six years. I can only hope that I helped make your time with me enjoyable and educational.

I am grateful for all of the professional development and social support I have received from my colleagues and friends here at Auburn. Thank you to all of the past and present members of the Geocognition Lab: Nick Soltis, Steph Courtney, Akilah Alwan, Lindsay

Maudlin, Haven Cashwell, Quinlan Swain, Tyler Smith, Ally Brown, Haylie Mikulak, and Hang Song. You have challenged and encouraged me to be my best self and I will carry your advice and wisdom with me wherever I go. I also would not have survived the COVID-19 pandemic without the friendship of the above as well as Morrighan O'Connell, Lucas Monroe, Connor Cain, Nora Lopez Rivera, and Pedro Montalvo. I am also grateful for the Side Track crew for being my office away from the office complete with unlimited chai, green tea, and hibiscus fizzes.

To my friends abroad, Julia Danis, Rowena Madavaram, DeJon Moore, Brittany Taylor, Biz Cross, and everyone else I've cried on the phone with: thank you! I know most of the time you have no idea what I'm talking about, but you've been there, nonetheless.

To my parents, Herbert and Michelle Johnson, and my brother, Zechariah Johnson: I feel the pride and love you have for me every day. Thank you for all you've done to help make me the person I am today.

Lastly, Damon, my partner and best friend, I think you know best of all what it has taken to get here. Thank you for always rooting for me. I love you.

## Table of Contents

Expanding the Measurement, Use, and Support of Spatial Reasoning in STEM and the Geosciences.....	1
Abstract .....	2
Acknowledgements.....	3
List of Tables .....	10
List of Figures .....	11
List of Abbreviations .....	12
Chapter 1: An Introduction for the Role and Importance of Exploring the Cognitive Domain in Geoscience Education Research .....	13
Objectives .....	14
Project 1: Measurement .....	15
Project 2: Support .....	15
Project 3: Use .....	16
References.....	17
Chapter 2: The Development and Analysis of a Survey for Exploring Trends in Spatial Reasoning Skills, Experiences, and Confidence in College-educated Adults .....	19
Abstract.....	20
Introduction.....	21
Methods and Materials.....	24

Recruitment and Participant Population .....	24
Survey Development and Validation .....	25
Survey Analysis .....	27
Results.....	27
SRCES Construct Validation.....	27
Discussion.....	29
High College to Career Consistency.....	30
Experience and Confidence Correlate to Spatial Skills .....	30
Non-STEM Outperforms STEM on Spatial Skills Test .....	31
Limitations .....	32
Conclusions.....	32
References.....	33
Chapter 3: Evaluating the Effectiveness of Digital and Physical Geologic-Block Models on Performance on a Structural Geology Activity and Spatial Assessment .....	43
Abstract.....	44
Introduction.....	45
Visual Aids in the Geosciences.....	46
Virtual Affordances for Spatial Learning in the Geosciences .....	47
Methods.....	49
Theoretical Framework.....	49

Participation and Recruitment.....	49
Activity Development and Description .....	50
Statistical Analyses .....	53
Results.....	54
Discussion.....	55
Effects on Spatial Visualization.....	55
Effects on Student Learning.....	56
Limitations .....	57
Conclusions.....	57
References.....	59
Chapter 4: Student Challenges with and Strategies for Solving Visuo-penetrative Structural	
Geology Tasks .....	70
Abstract.....	71
Introduction.....	72
Characterizing Visuo-penetrative Skills .....	72
Methods.....	75
Theoretical Framework.....	75
Participation and Recruitment.....	76
Analysis.....	77
Eye-tracking.....	77

Interview Data.....	79
Results.....	79
Heat Maps of Representative Items .....	80
Insights from Interview Data .....	83
Discussion.....	85
Students' Mental Visualizations .....	85
Recommendations.....	86
Limitations .....	87
Conclusion .....	87
Acknowledgements.....	88
Funding .....	88
Disclosure Statement .....	88
References.....	89
Chapter 5.....	106
Geospatial Thinking: Conclusions and Reflections on Spatial Thinking in the Geosciences ....	106
Measurement.....	107
Support.....	107
Use .....	108
Supplementary Materials A. Spatial Reasoning Confidence and Experience Survey (SRCES)	
Questions.....	109

Supplementary Materials B. Geologic Block Model Activity (No Block Model) ..... 112

Supplementary Materials C. Geologic Block Model Activity (Digital Block Model) ..... 119

Supplementary Materials D. Geologic Block Model Activity (Physical Block Model) ..... 126

## List of Tables

Demographic variables for the survey participants recruited from MTURK (n =1609).....	39
Exploratory factor analysis loadings for the Spatial Experience and Confidence questions.....	40
Confirmatory factor analysis for the Spatial Experience and Confidence questions.....	41
Demographic variables for the study participants (n=75).....	64
Exploratory factor analysis on the 16-item GBCT.....	65
Correct and incorrect responses separated by item and by error type.....	94
Codebook showing the codes created, the number of code applications, the description of the code and an example excerpt for each.....	103

## List of Figures

Physical, digital, and control visual aids given to participants with the activity.....	66
Violin plot showing the distribution of score changes from pre- to post-GBCT.....	67
Violin plot displaying student performance on a structural geology activity.....	68
Three examples of common errors that are made when solving for the cross-section in the block model.....	92
Histogram showing the distribution of scores on the GBCT.....	93
Heat maps of the item #1 from the GBCT to represent dipping beds.....	95
Heat maps of the item #4 from the GBCT to represent upright folds.....	96
Heat maps of the item #6 from the GBCT to represent basins.....	97
Heat maps of the item #7 from the GBCT to represent plunging folds.....	98
Heat maps of the item #15 from the GBCT to represent faulting beds.....	99
Code co-occurrences grid displaying the key elements of the block model and themes.....	100
Example solution for a geologic block model challenge.....	102

## List of Abbreviations

ANOVA	Analysis of Variance
CFA	Confirmatory Factor Analysis
DBER	Discipline-based Education Research
DEI	Diversity, Equity, and Inclusion
EFA	Exploratory Factor Analysis
GBCT	Geologic Block Cross-sectioning Test
GER	Geoscience Education Research
MTURK	Amazon Mechanical Turk
STEM	Science, Technology, Engineering, and Math
SRCES	Spatial Reasoning Confidence and Experience Survey
SRI	Spatial Reasoning Instrument

## Chapter 1

# **An Introduction for the Role and Importance of Exploring the Cognitive Domain in Geoscience Education Research**

Geoscience Education Research (GER) is a field of Discipline-Based Education Research (DBER) aimed at developing knowledge, skills, and behaviors that promote geoscience learning using scholarly and evidence-based research (St. John & McNeal, 2017). It encompasses spatial thinking, climate change communication, and diversity, equity, and inclusion (DEI), to name a few (St. John, et al., 2021). The studies described in this dissertation fall within the cognitive domain of GER which involves the mental processes that affect learning in the geosciences (Kastens et al., 2018). In fact, the research herein addresses each of the three grand challenges posed by Ryker et al. (2018): the use, measurement, and support of spatial thinking skills. These grand challenges will be expanded on in the next section and in the following chapters.

Spatial thinking can be defined as the skill of a person to conceptualize, represent, and process spatial information (National Research Council, 2006). Uttal et al. (2013b) defined spatial thinking as “the mental process of representing, analyzing, and drawing inferences from spatial relations...between objects or relations within objects” (p. 367). The strength of a person’s spatial thinking skills and their exposure to development opportunities through educational training can promote or inhibit their decision to select a STEM field of study (Wai et al., 2009). Wai et al. (2009) compiled 50 years of psychometric data to conclude that spatial ability is correlated with STEM achievement and subsequent career paths. Forty-thousand random participants were tracked for over 11 years to assess their self-organization into careers based on mathematical, verbal, and spatial ability. The authors show that students with high

spatial reasoning excelled in physical science, math/computer science, and engineering in terminal bachelor, master, and doctorate degrees. Students who had underdeveloped spatial skills tended to self-organize into education, law, and business. Occupations pursued after college strongly resemble those same trends.

This proposed dissertation work is based on evidence that spatial reasoning is malleable and can improve with various active or passive training opportunities (Uttal et al., 2013a). The malleability of spatial reasoning is a primary consideration when developing projects designed to promote student participation and success in STEM since spatial reasoning can predict student performance when addressed early on (Uttal & Cohen, 2012). Additionally, geoscientists routinely perform tasks that are inherently spatial. A review of the geocognition literature by Kastens and Ishikawa (2006) describes common tasks that can be grouped into either describing, comprehending, or metaphorically using spatial information. The first group includes recognizing patterns which can be used to classify minerals with similar crystal habit. The second group includes understanding the orientation, position, and processing of geological information, such as a fold structures seen on a topographic map and in profile view. The third group involves the transfer of spatial thinking skills to circumstances that are not inherently spatial, such as using a line to represent geologic time.

### **Objectives**

The aims of this dissertation are to address each of the grand challenges within geoscience education research on spatial reasoning: measuring spatial skills, supporting the development of spatial skills, and understanding the use of spatial skills (Ryker et al., 2018). I aim to further elucidate the relationship between spatial skills in academic settings and career choices. Then I will zoom in on the deployment of spatial skills in a geoscience-specific context.

## **Project 1: Measurement**

Measurement involves how to accurately quantify peoples' spatial skills, their beliefs about and experience with their spatial skills, and the relationship between those factors. In this project, I measured the spatial reasoning, spatial confidence, and spatial experience of the college-educated adult public using a novel instrument. The spatial reasoning skills were measured using a previously validated survey consisting of multiple defined spatial skills. Novel questions were created and validated with experts to assess spatial confidence and spatial experience. I define spatial confidence as the belief in the strength of one's own spatial thinking skills. I define spatial experience as opportunities one has had to use spatial reasoning. Other demographic variables that will be collected are race/ethnicity, gender, STEM or non-STEM college major, and STEM or non-STEM career field. I explored the relationship between measured spatial skills, spatial experience, spatial confidence, major, and careers. This project builds on other survey studies that sought to explore the relationship between STEM, spatial thinking, and career outcomes by providing another robust instrument for assessing other aspects of spatial learning outcomes.

## **Project 2: Support**

Support focuses on how GER can enhance the development of spatial skills through explicit geoscience instruction. In this project, we explored the impacts of a pedagogical tool, 3D geologic block model, on performance on a penetrative-thinking instrument. Penetrative thinking is a skill used in the geosciences to imagine what the interior of an object looks like given its exterior appearance. Participants used either a virtual, physical wooden, or no geologic block model. The block models were paired with a topographic- and geologic-map activity. In this activity, participants used whichever tool is given as an aid to visualize the subsurface geology, a

common geological task. Pre- and post-activity penetrative thinking scores were compared to determine if and how much the pedagogical tool type impacts penetrative thinking. This project emphasizes the need to understand the effectiveness of pedagogical tools on applicable spatial-skill development in students.

### **Project 3: Use**

Use involves the specific skills that are relevant in the geosciences and how students are actually employing them to complete geoscience tasks. I interpret this challenge to also include the process of problem solving using specific strategies. For example, some participants in prior studies have used logical reasoning algorithms to solve spatial problems without spatial visualization (Stull et al., 2012). This project involved using qualitative methods to explore and document the challenges that students report having on spatial tasks and what strategies they employ to overcome those challenges. The spatial tasks the participants were given are related to penetrative thinking, but the range of strategies they use could be broad. Eye-tracking, a tool used to acquire a wide range of attention metrics, along with retrospective interviews were used to gather detailed insight into the process of solving a geoscience-specific spatial task. This project emphasizes the need for more process-oriented inquiry into how these skills can be more efficiently developed, especially in geoscience novices.

## References

- Kastens, K. A., & Ishikawa, T. (2006). Spatial thinking in the geosciences and cognitive sciences: A cross-disciplinary look at the intersection of the two fields. *Special Papers-Geological Society of America*, 413, 53.
- Kastens, K. A., Dere, A., Pennington, D., & Richezza, V. (2018). "Research on Cognitive Domain in Geoscience Learning: Quantitative Reasoning, Problem Solving, and Use of Models". In St. John, K (Ed.) (2018). *Community Framework for Geoscience Education Research*. National Association of Geoscience Teachers.  
[https://doi.org/10.25885/ger\\_framework/8](https://doi.org/10.25885/ger_framework/8)
- National Research Council. (2006). Committee on support for thinking spatially.(2006). Learning to think spatially: GIS as a support system in the K-12 curriculum.
- Ryker, K., Jaeger, A. J., Brande, S., Guereque, M., Libarkin, J., & Shipley, T. F. (2018). "Research on Cognitive Domain in Geoscience Learning: Temporal and Spatial Reasoning". In St. John, K (Ed.) (2018). *Community Framework for Geoscience Education Research*. National Association of Geoscience Teachers.  
[https://doi.org/10.25885/ger\\_framework/7](https://doi.org/10.25885/ger_framework/7)
- St. John, K., & McNeal, K. S. (2017). The strength of evidence pyramid: One approach for characterizing the strength of evidence of geoscience education research (GER) community claims. *Journal of Geoscience Education*, 65(4), 363-372. doi: 10.5408/17-264.1
- St. John, K., McNeal, K. S., MacDonald, R. H., Kastens, K. A., Bitting, K. S., Cervato, C., . . . Teasdale, R. (2021). A community framework for geoscience education research:

- Summary and recommendations for future research priorities. *Journal of Geoscience Education*, 69(1), 2-13. doi:10.1080/10899995.2020.1779569
- Stull, A. T., Hegarty, M., Dixon, B., & Stieff, M. (2012). Representational translation with concrete models in organic chemistry. *Cognition and Instruction*, 30(4), 404-434.
- Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why, and how?. In *Psychology of Learning and Motivation*, 57,147-181). Academic Press.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013a). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139(2), 352–402. <https://doi.org/10.1037/a0028446>
- Uttal, D. H., Miller, D. I., & Newcombe, N. S. (2013b). Exploring and enhancing spatial thinking: Links to achievement in science, technology, engineering, and mathematics?. *Current Directions in Psychological Science*, 22(5), 367-373.
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101(4), 817-835. doi:10.1037/a0016127

## **Chapter 2**

# **The Development and Analysis of a Survey for Exploring Trends in Spatial Reasoning Skills, Experiences, and Confidence in College-educated Adults**

## **Abstract**

Recently, studies have investigated individuals' confidence in and experience with spatial skills as predictors for spatial skills test performance and STEM success. However, there is opportunity for more exploration into these affective factors, spatial skills, and college to career consistency since much of the literature focuses on students in academic settings. The scope of this study was to develop a spatial experience and confidence survey (SRCES) in combination with a spatial skills instrument to explore how STEM participation at the college and career levels impact measured spatial skills, as well as how individuals view their spatial thinking and their self-reported experience with spatial tasks. Results from this study showed that STEM participants had higher experience and confidence scores compared to non-STEM participants. However, non-STEM participants performed better on the spatial skills assessment used in this study than STEM participants. Despite these measured scores, there was high college to career consistency, which suggests that there are more factors at play after formal education that shape the development of spatial skills. Future work should investigate other latent determinants of STEM success and the role that spatial thinking plays beyond k-12 and college education.

## Introduction

Spatial reasoning is the cognitive skill of considering objects in space (Kastens & Ishikawa, 2006) and is linked to student success in Science, Technology, Engineering, and Math (STEM) fields. A longitudinal study followed students from high school through their occupations and found that students in STEM showed higher spatial reasoning skills than their non-STEM counterparts at all levels (high school, college majors, and occupation) (Shea et al., 2001; Wai et al., 2009). The measurement of spatial reasoning skills and STEM participation and success has been well documented in K-12 and undergraduate student populations (Shea et al., 2001; Taylor & Hutton, 2013; Wai et al., 2009; Wolfgang et al., 2003). Many studies have found that there are significant gender differences in spatial thinking skills with men outperforming women (see meta-analysis by Lauer & Yang, 2019). However, we know from the literature that spatial skills are malleable and can be improved with targeted training (Uttal et al., 2013), and this training can eliminate the gender gap in as little as one semester (Reynolds et al., 2006). Furthermore, spatial skill gender gaps can be accounted for by the differences in prior experience that participants have had with spatial tasks (Gold et al., 2018), indicating that these gaps are circumstantial rather than biological in origin.

Spatial reasoning measures include contextualized tests that focus on a skill or suite of skills explicitly tied to a discipline-specific task, like the Topographic Map Assessment (TMA) (Newcombe et al., 2015) or the Geologic Block Cross-Sectioning Test (GBCT) (Ormand et al., 2014). Other tests are decontextualized and focus on an individual skill, like mental rotation of 3D cube figures as seen in Peters and Battista (2008) or spatial orientation from Hegarty and Waller (2004). One such decontextualized instrument for assessing spatial skills is the Spatial Reasoning Instrument (SRI) (Ramful et al., 2017) which was validated using middle school-age

children and measures three components of spatial reasoning: mental rotation, spatial orientation, and spatial visualization. However, not many studies explore the spatial reasoning skills of the general adult population, although spatial skills have been posited to be more useful in this increasingly technological world (Bednarz & Bednarz, 2008; Goodchild & Janelle, 2010; Hespanha et al., 2009).

Recently, studies have linked participants' confidence in their spatial reasoning and their experience with spatial content to performance in STEM fields. In a study by McNeal et al. (2020), a regression showed that there were two factors that significantly predicted undergraduate students' performance on the Topographic Maps Assessment (TMA). These factors were students' performance on a different spatial assessment and their self-reported knowledge of topographic maps, which we interpret as a measure of spatial skill confidence in this context. Both factors were positively correlated with TMA performance which provides evidence of the connection between both spatial reasoning and confidence on STEM performance. Another study by Ferguson et al. (2015) showed that women teachers who have higher anxiety in their spatial reasoning skills experienced higher math anxiety. Furthermore, a study by Burte et al. (2020) found that spatially confident teachers are more likely to incorporate spatial reasoning in math teaching. Shin et al. (2016) compared the spatial reasoning skills and attitudes of elementary and secondary (social studies) education and geography majors. Spatial skills were measured using the Spatial Skills Test (SST) (Lee, 2005) which was designed to involve skills relevant to map reading, a skill useful in (and sometimes specific to) geography. Attitudes were measured using the Attitude toward Spatial Thinking Inventory (ATSTI; Shin et al., 2016) to gather data on habits and attitudes of participants. The geography majors scored higher on average on both the SST and ATSTI than the elementary and secondary education

majors, with elementary education majors scoring the lowest. While these studies collectively focus on spatial reasoning, anxiety, attitudes, confidence and experience with undergraduates and pre-service teachers, there exists an opportunity to expand to different populations and careers.

This study separates individuals into STEM and non-STEM majors and careers based on the expanded econometric theoretical framework generated by Perna (2004) and adopted by Xu (2013). This framework details that there are certain costs and benefits to career selection and trajectory that are related and unrelated to financial outcomes. Xu (2013) aligns these monetary and nonmonetary considerations of Perna (2004) with the intrinsic and extrinsic considerations of Seibert et al. (2001) to characterize variables that influence individual choices. An example of a monetary and extrinsic consideration is that STEM major and career fields have higher earning potential than their non-STEM counterparts (Melguizo & Wolniak, 2012), so individuals may self-select based on the weight of these factors.

In the self-selection phase in undergraduate education, it is known that the development of cognitive skills, like spatial reasoning, has an influence on a students' organization into STEM or non-STEM fields (Kell & Lubinski, 2013; Wai et al., 2009). Under-developed skills could result in a shift to a major where a student's current developed cognitive skills are better suited. For example, highly developed mathematical reasoning skills might result in better performance in math courses. These cognitive skills can be characterized as non-monetary and intrinsic which is what we seek to explore in this study. The research questions are as follows:

- (1) What is the relationship between college major and career?
- (2) What is the relationship between measured spatial skills, self-reported experience with spatial tasks, and self-reported confidence in spatial reasoning?

(3) Do STEM and non-STEM individuals differ significantly in spatial skills, experience, and confidence?

The Spatial Reasoning Confidence and Experience Survey (SRCES) was developed in this study to address these research questions.

### **Methods and Materials**

This study was designed to gather data on the college-educated public's spatial skill level, experience with spatial tasks, college major, and career field to better understand the relationships between these variables.

#### **Recruitment and Participant Population**

University Internal Review Board (IRB) approval was attained before this study began. Participants were recruited in two pulses, May 2020 and January 2021, from Amazon Mechanical Turk (MTURK), an online crowdsourcing system where participants can complete tasks, like surveys, for rewards (Shank, 2016). MTURK is a useful tool for a study of this scope because its samples' response quality is representative of traditional subject pools for psychology and/or social sciences (Paolacci et al., 2010, Paolacci & Chandler, 2014). The requirements for participants were adults over the age of 18 located within the United States in good standing as a MTURK task completer. MTURK participants were directed to the Auburn Qualtrics site to anonymously complete the survey where several additional checks were used to ensure quality participation both internally (attention checks) and externally (reporting survey ID to MTURK). If participants did not pass these checks, their responses were removed from the dataset. Furthermore, to explore the major to career progression, only respondents who went to a post-secondary institution and were currently employed were included in this study (N = 1609). The

authors did not have access to information that could identify individual participants during or after data collection.

### **Survey Development and Validation**

The survey was developed with the goal of assessing the spatial reasoning skills, confidence, and experiences of the college-educated US population, along with gathering demographic data like age, race/ethnicity, gender, college major, and career field. The survey consisted of two parts, the SRI and the SRCES, and each served a different purpose in this study. The SRI portion measured spatial skills, while the SRCES portion assessed spatial experiences and confidence of participants. All computational analyses were performed in R (R Development Core Team, 2014) and Jamovi for R (The jamovi project, 2021).

The SRI from Ramful et al. (2017) is a 30-question assessment that contains mental rotation, spatial orientation, and spatial visualization items. Mental rotation skill is addressed as rotating an object either clockwise or counterclockwise. Spatial orientation is described as viewing oneself in a space, map reading, and understanding cardinal and relative directions. Spatial visualization items include aspects of patterns, symmetry, and relationships between and within objects. There are 10 questions designed to measure each skill in the assessment. This instrument differs from the spatial composite described by Wai et al. (2009) because the validated SRI includes elements of spatial orientation and spatial visualization not substantially addressed in the spatial composite. Since the SRI was intended for use with youth aged 11-13, the reliability of the SRI was reexamined with the adult participants in this study to ensure that it was consistent using Cronbach's alpha. Omega was computed to satisfy more rigorous reliability measures (Peters, 2018).

The SRCES portion consisted of spatial confidence and spatial experience constructs (Supplementary Materials A). In this context, spatial confidence is interpreted as how a person feels about their performance on a spatial question, test, or task. To assess the participants' spatial confidence, three confidence questions for the overall assessment and six confidence questions in performing spatial tasks in daily life were included on a four-point Likert scale (4 = Strongly agree, 3 = Somewhat agree, 2 = Somewhat disagree, 1 = Strongly disagree). Spatial experience refers to how often participants are engaging with spatially rigorous tasks in their daily lives. Similarly, participant experience with spatial reasoning skills was collected through nine items on a five-point Likert scale (5 = Always, 4 = Frequently, 3 = Sometimes, 2 = Rarely, 1 = Never). The eighteen items were created by considering contemporary spatially-rigorous experiences, and they were qualitatively assessed and iteratively edited for face validity by a panel of expert researchers. Factor analysis was used to determine the covariances in the survey items where items that share much of the same variance can be said to be a part of the same factor (Knetka et al., 2019). The 1609 participants were split roughly in half, and an exploratory factor analysis (EFA) was used on the first half ( $n = 805$ ) while a confirmatory factor analysis (CFA) was used on the second half ( $n = 804$ ) as recommended by Bandalos and Finney (2010). According to Knetka et al. (2019), EFA is useful for identifying underlying patterns in the data and can help identify items that can be removed. The CFA on the other hand confirms the theoretical model generated in the EFA. After the EFA was performed, any nonconformant questions on SRCES were removed before the CFA was performed. Additionally, each portion of the survey (SRI, SRI confidence questions, spatial confidence, and spatial experience) as well as the overall survey was tested for reliability using Cronbach's alpha and omega.

## **Survey Analysis**

After these validity and reliability measures were performed, the data were analyzed to answer the research questions of this study. To answer the first question, participants were grouped based on whether they continued in a career similar to their major and descriptive statistics were used to compare them. The groups were STEM major and STEM career (SS), STEM major and non-STEM career (SN), non-STEM major and STEM career (NS), and non-STEM major and non-STEM career (NN). For the second question, a correlation matrix was produced with confidence, experience, and SRI score being the selected variables. A correlation is considered significant if the p-value is at or below 0.05. The third question was addressed by performing an independent samples t-test where experience, confidence, and measured spatial skills are the dependent variables and career was the independent variable. Cohen's *d* was included as a measure of evaluating the effect size of the t-tests. Career fields were split into STEM and non-STEM based on the National Science Foundation's delineation (National Science Foundation, n.d.) and have been agreed upon by a panel of interdisciplinary higher education professors.

## **Results**

The survey was administered to the adult US-based population recruited through Amazon MTURK. Table 1 shows the demographic distribution of the study participants. Their ages ranged from 18-79 ( $M = 37.8$ ,  $SD = 10.9$ ,  $Med = 35$ ).

### **SRCES Construct Validation**

Confidence and Experience items generated for this survey were analyzed using factor analysis to determine if the items could be grouped into subscales. An exploratory factor analysis (EFA) using two factors was performed on the initial 18 items (9 Confidence items and 9

Experience items) for half of the dataset ( $n = 805$ ) (Table 2). The two factors, Confidence and Experience, explained 21.3 and 20.0 percent of the variance respectively for a cumulative 41.3 percent. Two items from Experience, SE3 and SE6, were removed because they did not meet a loading threshold of 0.4 after the EFA. A confirmatory factor analysis (CFA) on the second half of the data ( $n = 804$ ) (Table 3) revealed that the Confidence items, SC4 and SC5, fell below the 0.4 threshold and was subsequently removed from the analysis. The CFA model's fit statistics after removing the aforementioned items are as follows: a CFI of 0.928, TLI of 0.913, SRMR of 0.0451, and RMSEA of 0.0703. The final CFA model's internal consistency was examined using Cronbach's alpha and McDonald's omega. The alphas and omegas were strong: 0.842 and 0.847 for Confidence (7 items) and 0.835 and 0.837 for Experience (7 items), respectively with an overall  $\alpha = 0.880$  and  $\omega = 0.887$ . The survey includes the 30-item SRI (for measuring spatial skills) with reliability scores of  $\alpha = 0.812$  and  $\omega = 0.815$ . Combined, the final 44 items in the SRCES included in the analysis was found to be reliable ( $\alpha = 0.715$ ,  $\omega = 0.744$ ).

Most of the participants who had STEM majors continued in a STEM career (75.3%). Similarly, most participants who had a non-STEM majors continued in a non-STEM career (86.1%). The participants who switched to a career field different than their major made up 18.7% of the sample (Table 1).

A Pearson correlation analysis was performed on the experience, confidence, and spatial skill score variables and showed that all relationships were significant. Confidence was positively correlated with experience (Pearson's  $r = 0.620$ ,  $p < 0.001$ ). Spatial skill score was negatively correlated with experience (Pearson's  $r = -0.501$ ,  $p < 0.001$ ) and confidence (Pearson's  $r = -0.159$ ,  $p < 0.001$ ).

T-tests revealed that experience, confidence, and spatial skill score were all significantly different between STEM and non-STEM careers. STEM participants ( $n = 652$ ) reported higher confidence scores ( $M = 3.23$ ,  $SD = 0.470$ ) than non-STEM participants ( $n = 936$ ,  $M = 3.07$ ,  $SD = 0.613$ ), and the difference was statistically significant ( $t_{(1592)} = -5.48$ ,  $p < 0.001$ , Cohen's  $d = -0.279$ ). Similarly, STEM participants ( $n = 652$ ) reported more experience ( $M = 3.43$ ,  $SD = 0.725$ ) than non-STEM participants ( $n = 932$ ,  $M = 3.07$ ,  $SD = 0.613$ ), and that difference was also significant ( $t_{(1582)} = -6.83$ ,  $p < 0.001$ , Cohen's  $d = -0.349$ ). Contrarily, non-STEM participants ( $n = 944$ ) performed better on the spatial skills test ( $M = 13.66$ ,  $SD = 5.589$ ) than the STEM participants ( $n = 665$ ,  $M = 12.72$ ,  $SD = 5.216$ ), and this result was statistically significant ( $t_{(1607)} = 3.43$ ,  $p < 0.001$ , Cohen's  $d = 0.174$ ).

## Discussion

While spatial reasoning is well-researched in formal education settings, there is a gap in our understanding of the general population and the patterns that exist comparing spatial reasoning amongst STEM and non-STEM careers. Furthermore, these studies tend to leave out the participants' confidence and experiences as a part of the spatial skill assessment and outcomes process. We sought to address both gaps in this study by developing a survey (SRCES) that met reasonable reliability and validity standards that could be used to evaluate spatial confidence and experience. From the results of this study, three major points have been identified. Firstly, the large majority of participants in this study persisted in the STEM or non-STEM career field in which they studied in school. Secondly, experience and confidence are significantly related to spatial skills in an unexpected way. Thirdly, STEM and non-STEM participants report expected experience and confidence scores, but non-STEM participants performed better on the spatial skills test than STEM participants.

## **High College to Career Consistency**

The consistency from college to career may be similar in number for STEM and non-STEM, but the reasons that can be attributed to the high number of participants pursuing a career similar to their college major vary for the field type. Xu (2012) reports through statistical analyses that STEM graduates tend to persist to a STEM career for monetary benefits, while non-STEM persist in non-STEM due to nonmonetary costs. While there are many factors for career consistency, Xu (2012) emphasizes that this consistency is largely dependent on academic performance. The literature overwhelmingly discusses the relationship between STEM achievement and spatial skills and that academic success, especially in STEM, can be predicted by spatial thinking skills (Uttal et al., 2013; Uttal & Cohen, 2012; Newcombe, 2017; Stieff & Uttal, 2015). For example, Sorby et al. (2018) discuss that targeted spatial training not only improved students' STEM course grades, but also improved STEM retention for women. Since spatial thinking can predict academic success in STEM, and academic success is linked to selecting a career consistent with major, it was expected to see higher spatial skill scores in the STEM career group. However, the study showed the opposite.

## **Experience and Confidence Correlate to Spatial Skills**

Confidence and experience were found to be positively correlated, which was expected. If one reports more experience with spatial thinking, it stands to reason they would also be more confident in the strength of those skills. However, spatial skills were negatively correlated to experience and confidence meaning the higher someone scored on the spatial skills test, the lower their reported confidence and experience. This result is contrary to what is reported in studies like Gagnier et al. (2022), where participants' spatial skill performance was positively correlated with their confidence and negatively correlated to their anxiety, though their study

sample was exclusively primary school teachers. Similarly, McNeal et al. (2020) found that undergraduate students' self-reported experience was positively correlated to a spatial skill activity. The reason for the unexpected outcome in the present study could be attributed to the rhetoric surrounding the spatial reasoning skills that people excelling in STEM are purported to have. For example, Muenks et al. (2020) discovered that even parents' beliefs about their students' skills influenced their decision to pursue STEM careers regardless of how developed their spatial skills are.

### **Non-STEM Outperforms STEM on Spatial Skills Test**

Despite these findings from other studies, we found that non-STEM participants significantly outperformed STEM participants on the spatial skills test. One potential explanation for this discrepancy is that after completing a program in a controlled college environment, many other factors are introduced that have the potential to increase, decrease, or stagnate spatial skill practice and development. For example, the farther someone gets in their degree and gain seniority in their profession, the more specific their tasks become which may require less practice of more general spatial skills tasks that were represented in the spatial skills test that was administered in the current study. Furthermore, the spatial skills test is a measure of mental rotation, spatial orientation, and spatial visualization which certainly are not the full scope of skills one could develop in college and beyond. However, it is expected that practiced spatial skills like one seen in the discipline or specific career should transfer to tasks that are not explicitly practiced (Wright et al., 2008). So, the question remains of what are non-STEM students doing after college that contributes to better spatial skill performance? More investigation is required to explore other dimensions of spatial skill training and the factors that contribute to college-career consistency including and beyond spatial skill development.

## **Limitations**

Although this survey was highly reliable, we recognize that it may not capture the full range of spatial skills, experience, and confidence questions. There are many spatial skill assessments, but a more general and practical one was chosen to reduce the need for prior content or context knowledge. Also, we rely on participants to self-report when answering experience and confidence questions which allows for uncertainty in data collection. The participants in this survey were slightly more educated than the U.S. population which means we could not be getting the full range of responses. However, a more educated sample is sufficient to explore the research questions in this study.

## **Conclusions**

Since there is a research gap in understanding individuals' confidence and experience about their spatial skills, this work can be regarded in the research community as a starting point for understanding the affective factors that may influence spatial skill development and college-career consistency. The scope of this study was to explore how STEM participation at the college and career levels could affect spatial skills, confidence in spatial reasoning, and experience with spatial tasks. The key findings of this study are that STEM participants had more experience with and higher confidence in spatial thinking skills, but non-STEM participants outperformed STEM in the spatial skills test that was administered. Future work should investigate other latent determinants of STEM success and the role that spatial thinking plays beyond k-12 and college education.

## References

- Akbay, L., Akbay, T., Osman, E. R. O. L., & Kiliç, M. (2019). Inadvertent use of ANOVA in educational research: ANOVA is not a surrogate for MANOVA. *Journal of Measurement and Evaluation in Education and Psychology*, *10*(3), 302-314.  
<https://doi.org/10.21031/epod.524511>
- Bandalos, D. L., & Finney, S. J. (2010). Exploratory and confirmatory factor analysis. *Quantitative Methods in the Social and Behavioral Sciences: A Guide for Researchers and Reviewers*. New York, Routledge.
- Bednarz, R. S., & Bednarz, S. W. (2008). The importance of spatial thinking in an uncertain world. In *Geospatial Technologies and Homeland Security* (pp. 315-330). Springer, Dordrecht.
- Burte, H., Gardony, A. L., Hutton, A., & Taylor, H. A. (2020). Elementary teachers' attitudes and beliefs about spatial thinking and mathematics. *Cognitive Research: Principles and Implications*, *5*(1), 1-18. <https://doi.org/10.1186/s41235-020-00221-w>
- Crocker, L., & Algina, J. (2008). Introduction to classical and modern test theory Mason. *OH: Cengage Learning*.
- Ferguson, A. M., Maloney, E. A., Fugelsang, J., & Risko, E. F. (2015). On the relation between math and spatial ability: The case of math anxiety. *Learning and Individual Differences*, *39*, 1-12. <https://doi.org/10.1016/j.lindif.2015.02.007>
- Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, *15*(6), 373-378. <https://doi.org/10.1111/j.0956-7976.2004.00687.x>
- Gold et al., 2018.....

- Goodchild, M. F., & Janelle, D. G. (2010). Toward critical spatial thinking in the social sciences and humanities. *GeoJournal*, 75(1), 3-13. <https://doi.org/10.1007/s10708-010-9340-3>
- Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32(2), 175-191. <https://doi.org/10.1016/j.intell.2003.12.001>
- Hespanha, S. R., Goodchild, F., & Janelle, D. G. (2009). Spatial thinking and technologies in the undergraduate social science classroom. *Journal of Geography in Higher Education*, 33(S1), S17-S27. <https://doi.org/10.1080/03098260903033998>
- Johnson, E. T., & McNeal, K. S. (2021). Student perspectives of the spatial thinking components embedded in a topographic map activity using an augmented-reality sandbox. *Journal of Geoscience Education*, 1-12. <https://doi.org/10.1080/10899995.2021.1969862>
- Kahane, L. H. (2007). *Regression basics*. Sage publications.
- Kastens, K. A., & Ishikawa, T. (2006). Spatial thinking in the geosciences and cognitive sciences: A cross-disciplinary look at the intersection of the two fields. *Special Papers-Geological Society of America*, 413, 53.
- Kell, H. J., & Lubinski, D. (2013). Spatial ability: A neglected talent in educational and occupational settings. *Roeper Review*, 35(4), 219-230.
- Khan, A., & Rayner, G. D. (2003). Robustness to non-normality of common tests for the many-sample location problem. *Advances in Decision Sciences*, 7(4), 187-206.
- Knekta, E., Runyon, C., & Eddy, S. (2019). One size doesn't fit all: Using factor analysis to gather validity evidence when using surveys in your research. *CBE—Life Sciences Education*, 18(1), rm1. <https://doi.org/10.1187/cbe.18-04-0064>
- Lowrie, T., Logan, T., & Ramful, A. (2017). Visuospatial training improves elementary students'

- mathematics performance. *British Journal of Educational Psychology*, 87(2), 170-186.  
<https://doi.org/10.1111/bjep.12142>
- McNeal, K. S., Ryker, K., Whitmeyer, S., Giorgis, S., Atkins, R., LaDue, N., Clark, C., Soltis, N. & Pingel, T. (2020). A multi-institutional study of inquiry-based lab activities using the Augmented Reality Sandbox: impacts on undergraduate student learning. *Journal of Geography in Higher Education*, 44(1), 85-107.  
<https://doi.org/10.1080/03098265.2019.1694875>
- Melguizo, T., & Wolniak, G. C. (2012). The earnings benefits of majoring in STEM fields among high achieving minority students. *Research in Higher Education*, 53(4), 383-405.  
<https://doi.org/10.1007/s11162-011-9238-z>
- Moè, A., Jansen, P., & Pietsch, S. (2018). Childhood preference for spatial toys. Gender differences and relationships with mental rotation in STEM and non-STEM students. *Learning and Individual Differences*, 68, 108-115.  
<https://doi.org/10.1016/j.lindif.2018.10.003>
- National Science Foundation. (n.d.). *Research areas*. NSF. Retrieved April 11, 2023, from [https://www.nsf.gov/about/research\\_areas.jsp](https://www.nsf.gov/about/research_areas.jsp)
- Newcombe, N. (2017), "Harnessing Spatial Thinking to Support Stem Learning", *OECD Education Working Papers*, No. 161, OECD Publishing, Paris, <https://doi.org/10.1787/7d5dcae6-en>.
- Newcombe, N. S., Weisberg, S. M., Atit, K., Jacovina, M. E., Ormand, C. J., & Shipley, T. F. (2015). The lay of the land: Sensing and representing topography. *Baltic International Yearbook of Cognition, Logic and Communication*, 10(1), 6.
- Nunnally, J. C. (1978). *Psychometric Theory* 2nd edition (New York: McGraw).

- Ormand, C. J., Manduca, C., Shipley, T. F., Tikoff, B., Harwood, C. L., Atit, K., & Boone, A. P. (2014). Evaluating geoscience students' spatial thinking skills in a multi-institutional classroom study. *Journal of Geoscience Education*, 62(1), 146-154.  
<https://doi.org/10.5408/13-027.1>
- Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, 23(3), 184-188.  
<https://doi.org/10.1177/0963721414531598>
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision making*, 5(5), 411-419.
- Perna, L. W. (2004). Understanding the decision to enroll in graduate school: Sex and racial/ethnic group differences. *The Journal of Higher Education*, 75(5), 487-527.  
<https://doi.org/10.1080/00221546.2004.11772335>
- Peters, G. J. (2018). The alpha and the omega of scale reliability and validity: why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *European Health Psychologist*, 16(2), 56-69.
- Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Perceptual and Motor Skills*, 105(3), 997-1014.  
<https://doi.org/10.2466/pms.105.3.997-1014>
- R Core Team (2021). *R: A Language and environment for statistical computing*. (Version 4.0) [Computer software]. Retrieved from <https://cran.r-project.org>. (R packages retrieved from MRAN snapshot 2021-04-01).

- Reynolds, S. J., Piburn, M., Leedy, D. E., McAuliffe, C. M, Birk, J. P., and Johnson, J. K. (2006). The Hidden Earth – Interactive, computer-based modules for geoscience learning, in Manduca, C. A. and Mogk, D. W., eds., *Earth and Mind: How Geologists Think and Learn about the Earth: Geological Society of America Special Paper 413*, pp. 157-170. [https://doi.org/10.1130/2006.2413\(12\)](https://doi.org/10.1130/2006.2413(12)).
- Ryker, K., Jaeger, A. J., Brande, S., Guereque, M., Libarkin, J., & Shipley, T. F. (2018). "Research on Cognitive Domain in Geoscience Learning: Temporal and Spatial Reasoning". In St. John, K (Ed.) (2018). *Community Framework for Geoscience Education Research*. National Association of Geoscience Teachers. [https://doi.org/10.25885/ger\\_framework/7](https://doi.org/10.25885/ger_framework/7)
- Seibert, S. E., Kraimer, M. L., & Liden, R. C. (2001). A social capital theory of career success. *Academy of Management Journal*, 44(2), 219–237. <https://doi.org/10.5465/3069452>
- Sheng, Y. (2008). Testing the assumptions of analysis of variance. *Best Practices in Quantitative Methods*, 324-340.
- Shank, D. B. (2016). Using crowdsourcing websites for sociological research: The case of Amazon Mechanical Turk. *The American Sociologist*, 47(1), 47-55. <https://doi.org/10.1007/s12108-015-9266-9>
- Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology*, 93(3), 604. <https://doi.org/10.1037/0022-0663.93.3.604>
- Shin, E. E., Milson, A. J., & Smith, T. J. (2016). Future teachers' spatial thinking skills and attitudes. *Journal of Geography*, 115(4), 139-146. <https://doi.org/10.1080/00221341.2015.1100654>

- Sorby, S., Veurink, N., & Streiner, S. (2018). Does spatial skills instruction improve STEM outcomes? The answer is 'yes'. *Learning and Individual Differences*, *67*, 209-222.
- Stieff, M., & Uttal, D. (2015). How much can spatial training improve STEM achievement?. *Educational Psychology Review*, *27*, 607-615.
- Taylor, H. A., & Hutton, A. (2013). Think3d!: Training spatial thinking fundamental to STEM education. *Cognition and Instruction*, *31*(4), 434-455.  
<https://doi.org/10.1080/07370008.2013.828727>
- The jamovi project (2021). *jamovi*. (Version 2.2) [Computer Software]. Retrieved from <https://www.jamovi.org>.
- Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why, and how?. In *Psychology of Learning and Motivation* (Vol. 57, pp. 147-181). Academic Press. <https://doi.org/10.1016/B978-0-12-394293-7.00004-2>
- Uttal, D. H., Miller, D. I., & Newcombe, N. S. (2013). Exploring and enhancing spatial thinking: Links to achievement in science, technology, engineering, and mathematics?. *Current Directions in Psychological Science*, *22*(5), 367-373.
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4), 817. <https://doi.org/10.1037/a0016127>
- Wolfgang, C., Stannard, L., & Jones, I. (2003). Advanced constructional play with LEGOs among preschoolers as a predictor of later school achievement in mathematics. *Early Child Development and Care*, *173*(5), 467-475.  
<https://doi.org/10.1080/0300443032000088212>
- Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S., & Kosslyn, S. M. (2008). Training

generalized spatial skills. *Psychonomic bulletin & review*, 15, 763-771.

Xu, Y. J. (2013). Career outcomes of STEM and non-STEM college graduates: Persistence in majored-field and influential factors in career choices. *Research in Higher Education*, 54(3), 349-382. <https://doi.org/10.1007/s11162-012-9275-2>

Table 1

*Demographic variables for the survey participants recruited from MTURK (n =1609).*

<b>Demographic</b>	<b>%</b>	<b>N</b>
<b>Race/Ethnicity</b>		
White/Caucasian	64.8	1042
Black/African/African American	19.4	312
Asian/Asian American	7	112
Latin American/Hispanic	5.2	84
American Indian/Native American	2.2	35
Other	.6	10
No Response	.6	9
Middle Eastern/North African	.2	3
Native Hawaiian/Pacific Islander	.1	2
<b>Gender</b>		
Male	59.5	958
Female	39.9	642
Other/Nonbinary/No Response	.6	9
<b>Highest Degree Earned</b>		
Bachelors	64.9	1045
Masters	26.5	427
Associates	6.8	109
Doctorate	1.7	28
<b>Career Field</b>		
STEM	41.3	665
Non-STEM	58.7	944
<b>College-Career Consistency</b>		
STEM to STEM	33.6	541
STEM to Non-STEM	11	177
Non-STEM to STEM	7.7	124
Non-STEM to Non-STEM	47.7	767

Table 2

*Exploratory factor analysis loadings for the Spatial Experience and Confidence questions. Factor loadings below 0.4 are not shown. Uniqueness is the portion of the variance not shared with any other variables. Note. 'Minimum residual' extraction method was used in combination with an 'oblimin' rotation.*

Item	Factor		Uniqueness
	1	2	
SC2	.706		.453
C2	.665		.531
SC1	.649		.454
SC5	.638		.668
SC3	.628		.562
C1	.620		.638
SC6	.603		.602
SC4	.504		.839
C3	.415		.534
SE5		.830	.398
SE4		.773	.450
SE7		.641	.509
SE8		.589	.540
SE1		.584	.494
SE9		.507	.672
SE2		.463	.630
SE6			.730
SE3			.859

Table 3

*Confirmatory factor analysis for the Spatial Experience and Confidence questions. SE is the standard error, Z is the z-score, and p is significance.*

<b>Factor</b>	<b>Item</b>	<b>Estimate</b>	<b>SE</b>	<b>Z</b>	<b>p</b>
Confidence	C2	.488	.0257	19.0	< .001
	SC1	.559	.0220	25.4	< .001
	C3	.583	.0300	19.4	< .001
	C1	.414	.0251	16.5	< .001
	SC3	.464	.0267	17.4	< .001
	SC6	.538	.0254	21.2	< .001
	SC2	.532	.0240	22.2	< .001
Experience	SE9	.485	.0350	13.9	< .001
	SE8	.788	.0395	20.0	< .001
	SE7	.775	.0403	19.2	< .001
	SE5	.923	.0417	22.2	< .001
	SE4	.837	.0403	19.2	< .001
	SE2	.524	.0309	17.0	< .001
	SE1	.786	.0341	23.1	< .001

## **Chapter 3**

# **Evaluating the Effectiveness of Digital and Physical Geologic-Block Models on Performance on a Structural Geology Activity and Spatial Assessment**

## **Abstract**

Digital representations are being used more frequently as pedagogical tools in classroom settings to support both spatial thinking skills and disciplinary content. One such tool that is used in geoscience instruction is geologic block models which can help students visualize complex subsurface morphologies, linking surface to subsurface relationships. Physical geologic block models can be difficult to design and acquire rendering them inaccessible for many practitioners. With the expansion of virtual learning spaces in recent years, it is becoming increasingly important to explore how digital versions of these physical visual aids impact student learning and spatial skill development. This study compares the efficacy of physical and digital geologic block models to not using a block model on student pre-post spatial skill assessment and performance during a novel structural geology activity. Results show that the students in the physical block model group improved marginally better than the digital block model and control groups on the spatial skill assessment, though that improvement was not significant. However, students in the digital block model group scored highest on the structural geology activity and had significantly higher activity scores than the control group. These results suggest that although visual aids in structural geology may not improve spatial thinking scores after a singular intervention, digital models have the potential to create improved learning outcomes for students. This finding broadens the possibilities for virtual learning efficacy and accessibility of pedagogical tools to more users.

## Introduction

Spatial visualization is a dimension of spatial thinking that involves many complex mental processes (Matthewson, 1999; Titus & Horsman 2009), but can be categorized into three components in the context of the geosciences: spatial relations, spatial manipulation, and visual penetrative skill (Titus & Horsman, 2009). Spatial relations includes the skill of mentally rotating an object around some axis (Shepard & Cooper, 1982). Spatial manipulation can include orientation where there is a mental transformation of an object into some other conformation or viewing it from a different perspective (Ekstrom et al., 1976; Titus & Horsman, 2009). Lastly is visual penetration which is the skill of imagining the interior of a solid object (Kali & Orion 1996). Visual penetrative skills, which we will use interchangeably with penetrative thinking skills, are not commonly included in spatial visualization categorizations but are important for many geoscientists' tasks to understand the formation history and morphology of the Earth (Titus & Horsman, 2009). For example, visual penetrative skills may be employed when a geoscientist is creating a cross-section from data that is gathered at the Earth's surface.

Titus and Horsman (2009) sought to determine whether geospatial practice in geology courses could improve students' visualization skills. The researchers lump the penetrative skill in with visualization due to its criticality in successfully performing geoscience tasks like creating and interpreting cross-sections. Two studies' findings were presented in this article. The first study involved administering pre- and post-assessments for spatial visualization to introductory, intermediate, and advanced geology courses. The second study qualitatively analyzed the students' performance on spatial visualization tasks before and after a structural geology course. The first study found that final performance in the course and spatial reasoning were linked where students with low spatial visualization skills were more likely to receive poor grades.

Furthermore, more spatially complex upper-level courses showed students' spatial visualization skills improved by the end of the courses. In the second study, performance on spatial visualization exercises showed that extra visualization practice improved student skills, but it was not necessarily indicative of better performance on exams. The utility of these exercises was that they showed when students were having trouble, but not necessarily how they were having trouble or why.

In the vein of geoscience courses contributing to spatial thinking improvement, Hannula (2019) developed a study designed to determine whether field geology courses can develop spatial thinking skills. One of these skills was penetrative thinking which was measured by the Geologic Block Cross-Sectioning Test (GBCT) (Ormand et al., 2014). Statistical analysis revealed that students' penetrative thinking skills improved from the start of the field methods class to the end. They also found that students' post-test scores on the GBCT remained the same from the end of the field methods class to the beginning of the next course in their sequence, showing that there is lasting effect of the training they received in the field methods class. The activities involved in this field methods course included mapping in the field and creating cross-sections using real-world landscapes as visualization aids.

### **Visual Aids in the Geosciences**

In a field as spatially complex as the geosciences, visual aids are widely used to help students understand the myriad processes and morphologies that exist. Departing from traditional modes of teaching, interactive visualizations can help bridge understanding gaps by moving from static displays to more engaging representations requiring direct interaction from students (Libarkin & Brick, 2002). Practitioners have often needed to get creative using materials like foam balls, construction paper, and washers to demonstrate and visualize concepts like Earth's

seasonality, faults, and half-life (Gray et al., 2011). Bateman et al. (2022) found that using playdough helped scaffold students' conceptual understanding of geologic structures in an introductory geoscience course.

More sophisticated interactive visual aids include the stream table which can simulate both spatial and temporal scales of fluvial processes (Cheek et al., 2017). A study by Lillquist and Kinner (2002) showed that when students worked with the interactive stream table, they had better exam scores and higher course evaluations than using traditional static representations. Clark and Linneman (2005) similarly saw higher motivation and interest in geomorphology research by engaging with labs developed around the stream table.

Another notable interactive visual aid used in the geosciences is the augmented-reality (AR) sandbox. The AR sandbox is a physical box of sand that has a projected image of a topographic map on the surface that updates in real time as the sand is manipulated (Woods et al., 2016). Though the AR sandbox has been lauded for its high potential in helping students understand topographical content better than its 2D map counterpart (Woods et al., 2016; Kundu et al., 2017; Nawaz et al., 2017), it has had mixed reviews on its efficacy (Giorgis et al., 2017; McNeal et al., 2020; Jackson et al., 2019). However, students generally are more engaged with the AR sandbox regardless of learning gains (Soltis et al., 2020; Johnson & McNeal, 2022).

### **Virtual Affordances for Spatial Learning in the Geosciences**

Bond and Cawood (2021) discuss that while virtual pedagogical tools and resources were an emergent field of inquiry for some time, there has been a rapid advancement in virtual teaching resources in recent years. For example, several studies show that by using virtual field trips, they were able to see gains in student content knowledge (Klippel et al., 2019) and student learning outcomes and experiences (Markowitz et al., 2018; Mead et al., 2019). Additionally,

participants in a study by Bond and Cawood (2021) performed better overall on a spatial thinking exercise when using the virtual outcrop model than by using a field photograph or geologic block models.

Research with block models and spatial thinking has been expanded into the virtual space by researchers like Reynolds et al. (2006) who have developed online modules designed to address student learning needs in geoscience courses. These online modules consisted of topography and 3D block models, the latter of which focused on the visual penetrative skill. The study consisted of one group that completed the virtual modules and one control group that did similar paper-based activities. Students were tested with two spatial reasoning tests and one geospatial reasoning test before and after a course. The researchers found that the students who used the virtual modules had higher gains than students from the control group, although both groups improved after completing the activities. The findings suggest that activities designed to be visually interactive can potentially improve spatial visualization skill, although this skill was not directly measured in this study by the selected spatial tests, but this skill was addressed substantially in the modules. Akciz and Thurmond (2018) also provide examples of these visual aids that have been used including physical geologic block models, like wooden or 3D-printed, and computer-based models.

Similarly, researchers have created tools, like *Visible Geology*, that allow students to generate their own geologic block models including adding geologic units, deformation events (like folding and faulting), and unconformities in a virtual 3D environment (Cockett, 2012; Cockett et al., 2016). The authors claimed that by implementing this tool with a large student group, they were able to observe more engagement with the material than usual with a lab exercise which allowed for more practice and interpretation with these visualization skills.

Over time, scientists and educators have found ways to help students visualize complex structural geology information. However, investigation into how these models compare to each other in efficacy is limited. According to Ryker et al. (2019), one of the three research priorities for spatial (and temporal) reasoning is to understand how different instructional methods and tools can simultaneously support content learning and spatial reasoning. To that end, we aim to explore student performance on a geospatial thinking assessment before and after using geologic blocks. Our research question is: How can using physical or digital geologic block models during a structural geology activity support student performance and improvement on a penetrative-thinking spatial assessment compared to not using any geologic block models?

## **Methods**

### **Theoretical Framework**

It is well-documented that spatial thinking broadly can improve with exposure and training (Uttal et al., 2013). Specifically, student spatial visualization skills have been shown to improve regardless of initial skillsets (e.g. Lord, 1987; Sorby, 2001). Orion and colleagues (1997) found that just by participating in a spatially intensive course, like geology, student visualization skills improve. We employ this finding in the development and execution of this study since geologic tasks are inherently spatial and discovering novel tools for increasing these skills along with appropriate coupling of educational strategy is a necessity (Libarkin & Brick, 2002).

### **Participation and Recruitment**

Participants were recruited from the undergraduate student population at a large-enrollment university in the Southeastern United States via email and classroom visits. They were targeted from a variety of academic disciplines and years. Internal Review Board (IRB)

approval was obtained from the participating university before the study commenced. Each student was consented to participate before and were compensated after completing the study. This study was conducted in a low-distraction laboratory space with the researcher where one student participated at a time. Once in the lab, students completed the GBCT pre-test to establish their initial penetrative thinking scores. Participants were then randomly assigned to one of three groups to complete the activity, with 25 participants in each group for a total of 75 participants. The first group used physical block models, the second used digital block models, and the control group did not use any block models as they only used the 2D map provided in the worksheets (Figure 1). The physical block model group was given wooden block models that correspond to the diagrams in the activity. The group assigned to digital models used the same block models, but they were accessed virtually via the website, Sketchfab, using the laboratory desktop computer. Both these physical models and their virtual counterparts were created by Ryan Hollister. The control group did not use any block model to complete the activity. After the students completed the activity, they took the GBCT post-test which is a version of the pre-test using the same geologic block representations but with different slicing planes.

### **Activity Development and Description**

The structural geology activity (Supplementary Materials A-C) was developed with researchers and practitioners in the geosciences and geoscience education research. Their practitioner knowledge and experiences in the classroom helped to identify where their students have historically struggled with understanding and interpreting geologic structures. Three broad categories of learning objectives were developed to encompass these challenges, and the activity was composed of three exercises with each corresponding to a learning objective. The student could earn up to 25 total points on the activity. It began with a brief overview of what geologic

and topographic maps are, what information they communicate, and their utility for geoscientists. A short video was created for each exercise to introduce the concepts in the activity and to simulate a brief lecture that usually would be given in a classroom laboratory. The learning objectives and exercises are as follows:

**1. Identify contour lines and sketch topographic profiles from elevation data.** Contour lines on a topographic map are designed so that all points that lie on the same contour line are the same elevation (Rapp et al., 2007). The topographic profile can depict the shape of the landscape given the numerical elevation values of the contour lines. A key component of understanding the block models and diagrams in the activity involve interpreting 3D structures from 2D information, which poses a challenge for many students (Gobert, 2005). The rationale for including this is to scaffold students toward the penetrative thinking exercises later.

In the first task of this activity, the students were presented a topographic map and asked to orient themselves on the map to determine where specific marked locations are relative to each other. Using the contours, the students needed to determine which locations were at the highest and lowest elevations. Then, the students were given a visibility challenge where they were asked if they can see a campfire at one location from another location on the map. Lastly for this topographic map, they were asked to select the most correct topographic profile for a line meant to represent a cross-section on the map.

The next map the students were given was a topographic map that included symbology for various locations (e.g., circle, triangle) and geologic unit ages (e.g., Pg, K). The students were asked to do another visibility challenge and to construct a topographic profile for a specified line meant to represent a cross-section on the map.

- 2. Identify and interpret the contacts between geologic units on geologic maps.** The patterns seen on geologic maps are the result of both the distribution of the rock units underground and the shape of the Earth's landscape at the surface. Therefore, a student needs to be able to read topographic information, identify the distribution of rocks at the surface, and activate spatial visualization skills to imagine the subsurface. The dip of geologic units is the direction that they descend into the ground from  $0^\circ$  to  $90^\circ$  from the horizontal (Bursztyn, 2021). Dip is also an important consideration for students trying to understand the structural geology being represented by a geologic map. Understanding dip is a key part of interpretation, especially determining what a cross-section of the map would look like including the geologic units. In this portion of the activity, the students were given two geologic maps that had topographic information, symbology for contacts (or the surfaces between the different rock units), and geologic unit ages. They were asked to describe the relationship between the topography and the contacts. Afterwards, they estimated how steeply and in what direction these geologic units dipped using the surficial information.
- 3. Determine the direction of dip and estimate the dip angle of geologic units and structures from geologic map data.** The third group of exercises in this activity builds on topographic information, contacts, and dips from the previous exercises and includes the concept of faults. Faults are planar breaks in the sequence of rock units that offset those units. The three maps the students were given added new symbology for fault lines. The students were asked to estimate the dip angle and determine dip direction of contacts and faults shown on the geologic maps. Additionally, the students determined the kind of

fault that was shown on the map (e.g., normal, reverse, or strike-slip) or, rather, the relative direction of motion or offset of the block portions.

The students were given the same activity sheets, except the sheets given to the physical block model group instructed them to trace their finger over the wooden blocks and the 2D map provided, the digital model group were instructed to trace their mouse across the digital model and the 2D map, and the control group was instructed to only trace their finger along the 2D map. To supplement any knowledge deficits for students who were unfamiliar with these concepts, a list of technical terms with definitions were included on a separate sheet of paper for the students to reference. The students needed to activate their penetrative thinking (or other spatial) skills and use the block models as visual aids to solve the questions therein.

### **Statistical Analyses**

Although the student population is similar to Ormand et al. (2014) an exploratory factor analysis was used to determine whether all of the items on the GBCT were loading on the same factor for this group of students. Items that fell below a threshold of 0.3 were removed from the analysis as they may not be great indicators of the penetrative thinking skill we hoped to test. A reliability analysis of the GBCT pre-test with this group of participants was also conducted to determine how consistently the GBCT was measuring penetrative thinking skills. Reliability was tested with the expectation of a Cronbach's alpha score of .70-.75 or better (Nunnally, 1978; Ponterotto and Ruckdeschel, 2007).

Items from the Spatial Reasoning Confidence and Experience Survey (SRCES) (Johnson & McNeal, in prep) were used in a linear regression model to determine if the amount of spatial experience a student has and their confidence in their spatial thinking influenced performance on

either the GBCT or the structural geology activity. To measure group differences in GBCT pre- and post-test scores, the analysis of variance (ANOVA) statistical test was used for the physical, virtual, and control block model groups. ANOVA is a useful statistical method for determining if there is a significant difference between groups on a dependent variable (Akabay et al., 2019). ANOVA was also used to determine if there was a significant difference between block model groups on activity score. The ANOVA was deemed appropriate since it is robust against certain normality violations (Blanca et al., 2017). All computational analyses were performed in R (R Development Core Team, 2014) or jamovi for R (The jamovi project, 2021).

## Results

An exploratory factor analysis on the GBCT pre-test showed 5 items did not make it above the 0.3 loading threshold and were subsequently removed from the analysis leaving 11 items. Reliability analysis of the 11-item GBCT produced a Cronbach's  $\alpha$  of 0.821 and McDonald's  $\omega$  of 0.824. Out of the 75 participants, 79.9% were white, 65.3% were women, 69.4% were first- and second-year students, and 54.7% were enrolled in a STEM major (Table 1).

Linear regression showed that experience and confidence did not significantly predict GBCT improvement ( $R^2 = 0.116$ ,  $F_{(2,42)} = 2.76$ ,  $p = 0.074$ ). Likewise, experience and confidence did not significantly predict performance on the activity ( $R^2 = 0.0818$ ,  $F_{(2,42)} = 1.87$ ,  $p = 0.167$ ).

Results of the ANOVA showed that there was not a significant difference in pre- and post-GBCT scores regardless of the block model group the participants were assigned to ( $F_{2,72} = 0.150$ ,  $p = 0.861$ ), although the physical group showed the most improvement ( $M = 0.520$ ,  $SD = 2.08$ ) over the digital ( $M = 0.160$ ,  $SD = 2.29$ ) and control ( $M = 0.360$ ,  $SD = 2.60$ ). However,

there was a significant difference in activity performance between block model groups ( $F_{2,72} = 3.61, p = 0.032, \eta^2 = 0.091, \omega^2 = 0.065$ ). Out of the 25 total points that could be earned on the activity, the digital block model group scored the highest ( $M = 19.8, SD = 3.98$ ), the physical group scored the next highest ( $M = 17.9, SD = 4.02$ ), and the control group scored the lowest ( $M = 17.1, SD = 3.11$ ) (Figure 2). A Tukey's post-hoc comparison revealed that only the digital and control block model groups performed significantly different on the activity ( $t_{72} = 2.617, p = 0.029$ , Cohen's  $d = 0.740$ ).

## Discussion

To determine the efficacy of various structural geology visual aids, several statistical analyses were performed to arrive at the following findings: block models did not significantly improve scores on a spatial visualization test, but digital block models showed significantly higher scores on the structural geology activity.

### Effects on Spatial Visualization

While the physical block model group did show more improvement than the digital block model and control groups, it was by less than half a point out of the 11-point GBCT. This result was not significant and there are several insights from the literature that could offer an explanation. The students may not have had enough time to practice with the structural geology activity and the blocks to show significant skill gains. For example, Gold et al. (2018) had 11 targeted spatial training modules during an academic semester which produced significant spatial thinking gains compared to a control. Similarly, Hannula (2019) found that student spatial thinking improved and gender gaps closed after a full semester of a geology field course. In an experiment by Cohen and Hegarty (2014) where a short intervention produced significant spatial thinking gains, their training methodology involved directly correcting participants' incorrect

attempts which could point to the need for real-time correction in the training process for one-off interventions.

Furthermore, the current study explores gains in one particular measure of spatial thinking, spatial visualization or penetrative thinking, that we expected to align with the structural geology activity. However, the literature suggests that students could be employing any variety of discrete skills to solve spatial challenges. Ormand et al. (2014) showed that spatial skills are not monolithic, so a pre-post-test model that attempts to isolate any one skill may not fully capture gains in others.

### **Effects on Student Learning**

Students using visual aids performed as expected on the structural geology activity where those groups that used the geologic block models outperformed the control group. The digital block model group significantly outperformed the control group. This result could be attributed to the difference between static vs. dynamic visualizations. Libarkin and Brick (2002) discuss that static visualizations, like the 2D maps and diagrams used in the control group, can visually limit the students. Using a more interactive visualization instead allows the student to explore and discover those 2D to 3D connections in real time. These findings mirror a study by Bond and Cawood (2021) that show that students performed better when using a 3D block diagram than using traditional 2D visualizations. Furthermore, the interactive components of the block models and the use of technology as an instructional tool allow for immediate feedback with the activity which likely aided in better conceptual models, as seen in McNeal et al. (2008).

In this study, it is possible that the above mentioned reasons demonstrate why the digital group outperforms the physical group, but an alternative explanation could also be simply that students are more proficient at navigating web-based environments than handling the physical

geologic block models. More investigation is needed for the gap in performance between the digital and physical block model groups and the lack of a significant performance gap between the physical block model group and the control.

### **Limitations**

Although steps were taken to simulate an instructional environment (e.g., lecture videos and activity sheets), the students were still participating individually and lacked social context. Alternatively, students may have been able to concentrate more without distractions that come with a lab classroom setting. Participants spent about 75 minutes on the study, with about 30 to 45 of those minutes working on the activity, which is shorter than a typical 2-hour introductory geology lab. Additionally, students were not particularly encouraged to interact with the researcher to ask questions or gain clarity which may have led to individual learning challenges not being addressed. Lastly, the sample sizes for each treatment group in this study are small compared to studies that are multiple lab sections (e.g., Giorgis et al., 2017), or multi-institutional (e.g., McNeal et al., 2020).

### **Conclusions**

There is great potential in visual aids for teaching geoscience content and addressing spatial thinking gaps in students and further investigation is needed into the efficacy of digital and physical representations. This study investigated how effective physical and digital block models were in improving student spatial visualization skill and student performance on a novel structural geology assignment. Our findings shows that while spatial visualization skills did not significantly improve after students completed the activity regardless of treatment group, the digital block model group outperformed both other treatment groups on the activity and significantly more so than the control group. This study highlights the high potential for digital

visualizations to support student learning and the opportunity for increased accessibility of geoscience pedagogical tools.

## References

- Akbay, L., Akbay, T., Osman, E. R. O. L., & Kiliñç, M. (2019). Inadvertent Use of ANOVA in Educational Research: ANOVA is not A Surrogate for MANOVA. *Journal of Measurement and Evaluation in Education and Psychology*, *10*(3), 302-314.
- Akciz, S. O., & Thurmond, J. (2018, December). 3-D Printed Geologic Block Models to Improve Spatial Thinking Skills in Earth Sciences Courses. In *AGU Fall Meeting Abstracts* (Vol. 2018, pp. ED23F-0969).
- Bateman, K., Ham, J., Shipley, T., Tikoff, B., Barshi, N., & Ormand, C. (2020). Playdough modeling in geological field work to support spatial skills. *Journal of Geoscience Education*, *71*(1), 43-57.
- Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option?. *Psicothema*, *29*(4), 552–557.  
<https://doi.org/10.7334/psicothema2016.383>
- Bond, C. E., & Cawood, A. J. (2021). A role for Virtual Outcrop Models in Blended Learning - improved 3D thinking, positive perceptions of learning. *Geoscience Communication*, *4*, 233-244.
- Bursztyn, N., Riegel, H., Sajjadi, P., Masters, B., Zhao, J., Huang, J., ... & Klippel, A. (2021, March). Fostering geological thinking through virtual strike and dip measurements. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 303-308). IEEE.
- Cheek, K. A., LaDue, N. D., & Shipley, T. F. (2017). Learning about spatial and temporal scale: Current research, psychological processes, and classroom implications. *Journal of Geoscience Education*, *65*(4), 455-472.

- Clark, D. H., & Linneman, S. R. (2005). Combining stream table experiments, high-tech particle analysis, and the web to help geomorphology students evaluate landform evolution. *Journal of Geoscience Education*, 53(1), 110-115.
- Cockett, R. (2012, December). Visible Geology-Interactive online geologic block modelling. In *AGU Fall Meeting Abstracts* (Vol. 2012, pp. ED51A-0872).
- Cockett, R., Moran, T., & Pidlisecky, A. (n.d.). Visible Geology. In 3-D Structural Interpretation (pp. 53–63). American Association of Petroleum Geologists.  
<https://doi.org/10.1306/13561985m1113671>
- Cohen, C. A., & Hegarty, M. (2014). Visualizing cross sections: Training spatial thinking using interactive animations and virtual objects. *Learning and Individual Differences*, 33, 63-71.
- Ekstrom, R. B., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests*, 1976. Educational Testing Service.
- Giorgis, S., Mahlen, N., & Anne, K. (2017). Instructor-led approach to integrating an augmented reality sandbox into a large-enrollment introductory geoscience course for nonmajors produces no gains. *Journal of Geoscience Education*, 65(3), 283-291.
- Gobert, J. (2005). The effects of different learning tasks on model-building in plate tectonics: Diagramming versus explaining. *Journal of Geoscience Education*, 53(4): 444–455.
- Gold, A. U., Pendergast, P. M., Ormand, C. J., Budd, D. A., & Mueller, K. J. (2018). Improving spatial thinking skills among undergraduate geology students through short online training exercises. *International Journal of Science Education*, 40(18), 2205-2225.
- Gray, K., Owens, K. D., Steer, D. N., McConnell, D. A., & Knight, C. C. (2011). An exploratory

- study using hands-on physical models in a large introductory, earth science classroom: student attitudes and lessons learned. *The Electronic Journal for Research in Science & Mathematics Education*, 15(2).
- Hannula, K. A. (2019). Do geology field courses improve penetrative thinking?. *Journal of Geoscience Education*, 67(2), 143-160.
- Jackson, D., Kaveh, H., Victoria, J., Walker, A., & Bursztyn, N. (2019). Integrating an augmented reality sandbox challenge activity into a large-enrollment introductory geoscience lab for nonmajors produces no learning gains. *Journal of Geoscience Education*, 67(3), 237-248.
- The jamovi project (2021). *jamovi*. (Version 2.2) [Computer Software]. Retrieved from <https://www.jamovi.org>.
- Johnson, E. T., & McNeal, K. S. (2022). Student perspectives of the spatial thinking components embedded in a topographic map activity using an augmented-reality sandbox. *Journal of Geoscience Education*, 70(1), 13-24.
- Johnson, E.T., & McNeal, K.S. (manuscript in review). *The development and analysis of a survey for assessing spatial reasoning skills, experience, and confidence within the general public*.
- Kahane, L. H. (2007). *Regression basics*. Sage publications.
- Kali, Y., & Orion, N. (1996). Spatial abilities of high-school students in the perception of geologic structures. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 33(4), 369-391.
- Klippel, A., Zhao, J., Oprean, D., Wallgrün, J. O., Stubbs, C., La Femina, P., & Jackson, K. L. (2020). The value of being there: Toward a science of immersive virtual field trips. *Virtual Reality*, 24(4), 753-770.
- Kundu, S. N., Muhammad, N., & Sattar, F. (2017, December). Using the augmented reality

- sandbox for advanced learning in geoscience education. In *2017 IEEE 6th international conference on teaching, assessment, and learning for engineering (TALE)* (pp. 13-17).  
IEEE.
- Libarkin, J. C., & Brick, C. (2002). Research methodologies in science education: Visualization and the geosciences. *Journal of Geoscience Education*, *50*(4), 449-455.
- Lillquist, K. D., & Kinner, P. W. (2002). Stream tables and watershed geomorphology education. *Journal of Geoscience Education*, *50*(5), 583-593.
- Lord, T. R. (1987). A look at spatial abilities in undergraduate women science majors. *Journal of Research in Science Teaching*, *24*(8), 757-767.
- Markowitz, D. M., Laha, R., Perone, B. P., Pea, R. D., & Bailenson, J. N. (2018). Immersive virtual reality field trips facilitate learning about climate change. *Frontiers in Psychology*, *9*, 2364.
- Mathewson, J. H. (1999). Visual-spatial thinking: An aspect of science overlooked by educators. *Science Education*, *83*(1), 33-54.
- McNeal, K. S., Miller, H. R., & Herbert, B. E. (2008). The effect of using inquiry and multiple representations on introductory geology students' conceptual model development of coastal eutrophication. *Journal of Geoscience Education*, *56*(3), 201-211.
- McNeal, K. S., Ryker, K., Whitmeyer, S., Giorgis, S., Atkins, R., LaDue, N., ... & Pingel, T. (2020). A multi-institutional study of inquiry-based lab activities using the Augmented Reality Sandbox: impacts on undergraduate student learning. *Journal of Geography in Higher Education*, *44*(1), 85-107.
- Mead, C., Buxner, S., Bruce, G., Taylor, W., Semken, S., & Anbar, A. D. (2019). Immersive,

- interactive virtual field trips promote science learning. *Journal of Geoscience Education*, 67(2), 131-142.
- Nawaz, M., Kundu, S. N., & Sattar, F. (2017). Augmented reality sandbox and constructivist approach for geoscience teaching and learning. *International Journal of Educational and Pedagogical Sciences*, 11(6), 1645-1648.
- Nunnally, J. C. (1978). *Psychometric Theory* 2nd edition (New York: McGraw).
- Orion, N., Ben-Chaim, D., & Kali, Y. (1997). Relationship between earth-science education and spatial visualization. *Journal of Geoscience Education*, 45(2), 129-132.
- Ormand, C. J., Manduca, C., Shipley, T. F., Tikoff, B., Harwood, C. L., Atit, K., & Boone, A. P. (2014). Evaluating geoscience students' spatial thinking skills in a multi-institutional classroom study. *Journal of Geoscience Education*, 62(1), 146-154.
- Peters, G. J. (2018). The alpha and the omega of scale reliability and validity: why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *European Health Psychologist*, 16(2), 56-69.
- Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures. *Perceptual and motor skills*, 105(3), 997-1014.
- R Development Core Team. (2014). *R: A language and environment for Statistical Computing*. Vienna, Austria. Retrieved from <http://www.r-project.org/>
- Rapp, D. N., Culpepper, S. A., Kirkby, K., & Morin, P. (2007). Fostering students' comprehension of topographic maps. *Journal of Geoscience Education*, 55(1), 5-16.
- Reynolds, S. J., Piburn, M. D., Leedy, D. E., McAuliffe, C. M., Birk, J. P., & Johnson, J. K.

- (2006). The Hidden Earth—Interactive, computer-based modules for geoscience learning. Geological Society of America Special Papers, 413, 157-170.
- Ryker, K., Jaeger, A. J., Brande, S., Guereque, M., Libarkin, J., & Shipley, T. F. (2018). "Research on Cognitive Domain in Geoscience Learning: Temporal and Spatial Reasoning". In St. John, K (Ed.) (2018). *Community Framework for Geoscience Education Research*. National Association of Geoscience Teachers.  
[https://doi.org/10.25885/ger\\_framework/7](https://doi.org/10.25885/ger_framework/7)
- Shepard, R. N. (6). Cooper. LA (1982). Mental Images and Their Transformations.
- Soltis, N. A., McNeal, K. S., Atkins, R. M., & Maudlin, L. C. (2020). A novel approach to measuring student engagement while using an augmented reality sandbox. *Journal of Geography in Higher Education*, 44(4), 512-531.
- Sorby, S. A. (2001). A course in spatial visualization and its impact on the retention of female engineering students. *Journal of women and minorities in science and engineering*, 7(2).
- Titus, S., & Horsman, E. (2009). Characterizing and improving spatial visualization skills. *Journal of Geoscience Education*, 57(4), 242-254.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139(2), 352–402. <https://doi.org/10.1037/a0028446>
- Woods, T. L., Reed, S., Hsi, S., Woods, J. A., & Woods, M. R. (2016). Pilot study using the augmented reality sandbox to teach topographic maps and surficial processes in introductory geology labs. *Journal of Geoscience Education*, 64(3), 199-214.

Table 1

*Demographic variables for the study participants (n =75).*

<b>Demographic</b>	<b>%</b>	<b>N</b>
<b>Race/Ethnicity</b>		
White/Caucasian	79.7	59
Asian/Asian American	6.8	5
Other	5.4	4
Black/African/African American	2.7	2
No Response	2.7	2
Latin American/Hispanic	1.4	1
American Indian/Native American	1.4	1
Middle Eastern/North African	1.4	1
<b>Gender</b>		
Female	65.3	49
Male	29.3	22
Nonbinary	2.7	2
No response	2.7	2
<b>Academic Year</b>		
First-year	38.7	29
Second-year	30.7	23
Third-year	17.3	13
Fourth-year or higher	13.3	10
<b>Major</b>		
<b>STEM</b>	<b>54.7</b>	<b>41</b>
<i>Science and Mathematics</i>	<i>36.0</i>	<i>27</i>
<i>Architecture, Design, and Engineering</i>	<i>10.7</i>	<i>8</i>
<i>Agriculture and Forestry</i>	<i>6.7</i>	<i>5</i>
<i>Zoology and Veterinary Medicine</i>	<i>1.3</i>	<i>1</i>
<b>Non-STEM</b>	<b>44.1</b>	<b>33</b>
<i>Arts and Humanities</i>	<i>14.7</i>	<i>11</i>
<i>Management, Business, and Finance</i>	<i>14.7</i>	<i>11</i>
<i>Education and Social Services</i>	<i>6.7</i>	<i>5</i>
<i>Health Care and Allied Health</i>	<i>2.7</i>	<i>2</i>
<i>Trades and Transportation</i>	<i>2.7</i>	<i>2</i>
<i>Hospitality, Tourism, and Service</i>	<i>1.3</i>	<i>1</i>
<i>Law, Law Enforcement, and Political Science</i>	<i>1.3</i>	<i>1</i>
No major identified	1.3	1

Table 2

*Exploratory factor analysis on the 16-item GBCT.*

	Factor	
	1	Uniqueness
pre4	0.7432	0.448
pre3	0.7276	0.471
pre2	0.6845	0.532
pre5	0.6098	0.628
pre8	0.5745	0.670
pre1	0.4873	0.763
pre15	0.4845	0.765
pre16	0.4624	0.786
pre9	0.4471	0.800
pre10	0.3930	0.846
pre12	0.3349	0.888
pre6	0.2920	0.915
pre13	0.2825	0.920
pre7	0.1534	0.976
pre11	0.1441	0.979
pre14	0.0414	0.998

Note. 'Minimum residual' extraction method was used in combination with a 'oblimin' rotation

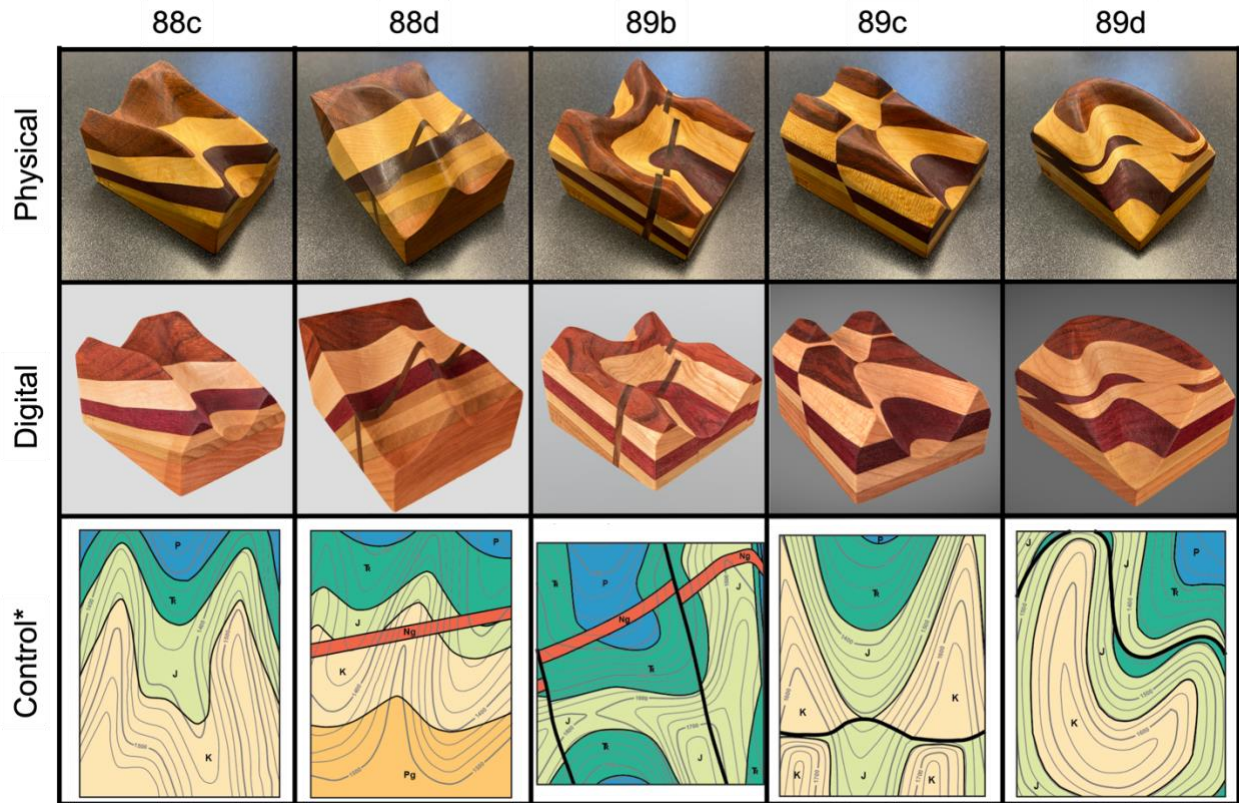


Figure 1. Physical, digital, and control visual aids given to participants with the activity. \*Every participant was given the 2D maps in the activity worksheets, but the control group only received the 2D maps.

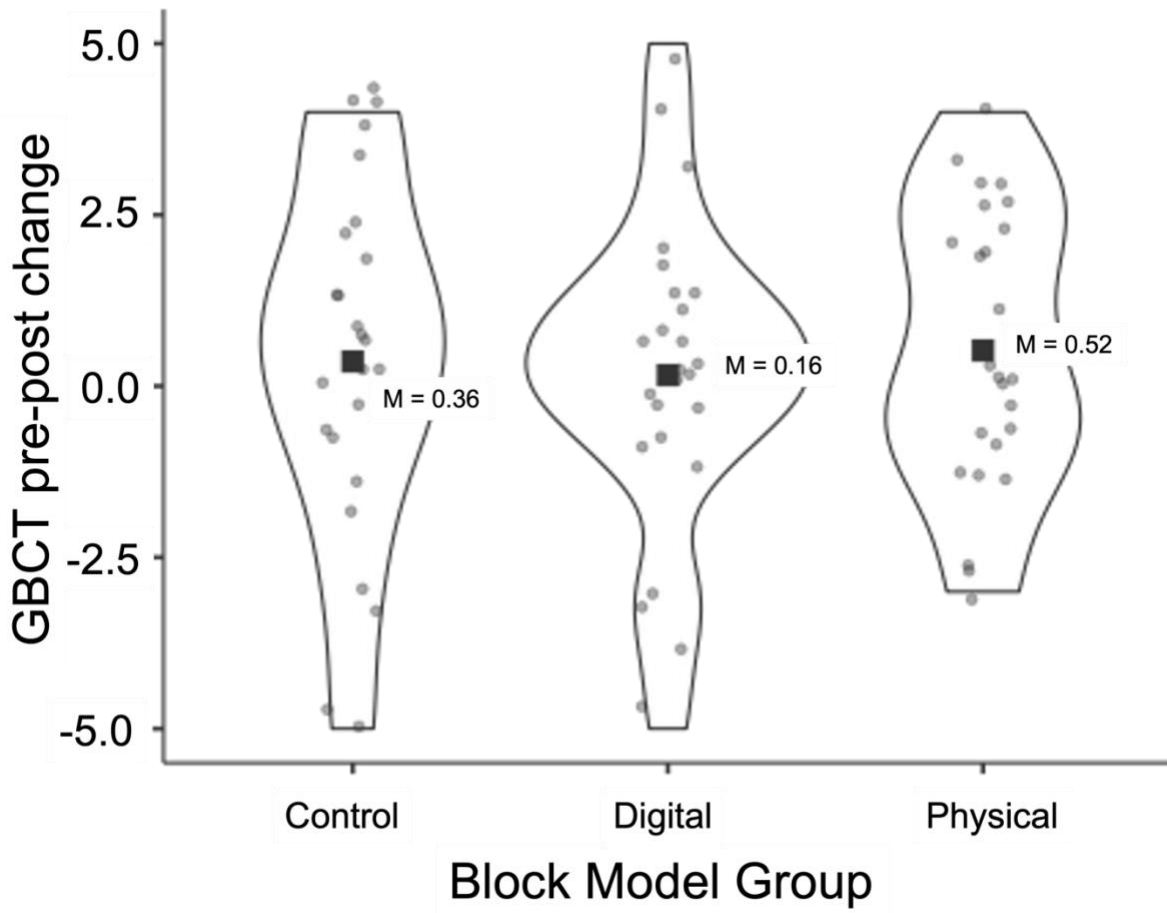


Figure 2. Violin plot showing the distribution of score changes from pre- to post-GBCT separated by block model group (control, digital, or physical). Black squares represent means.

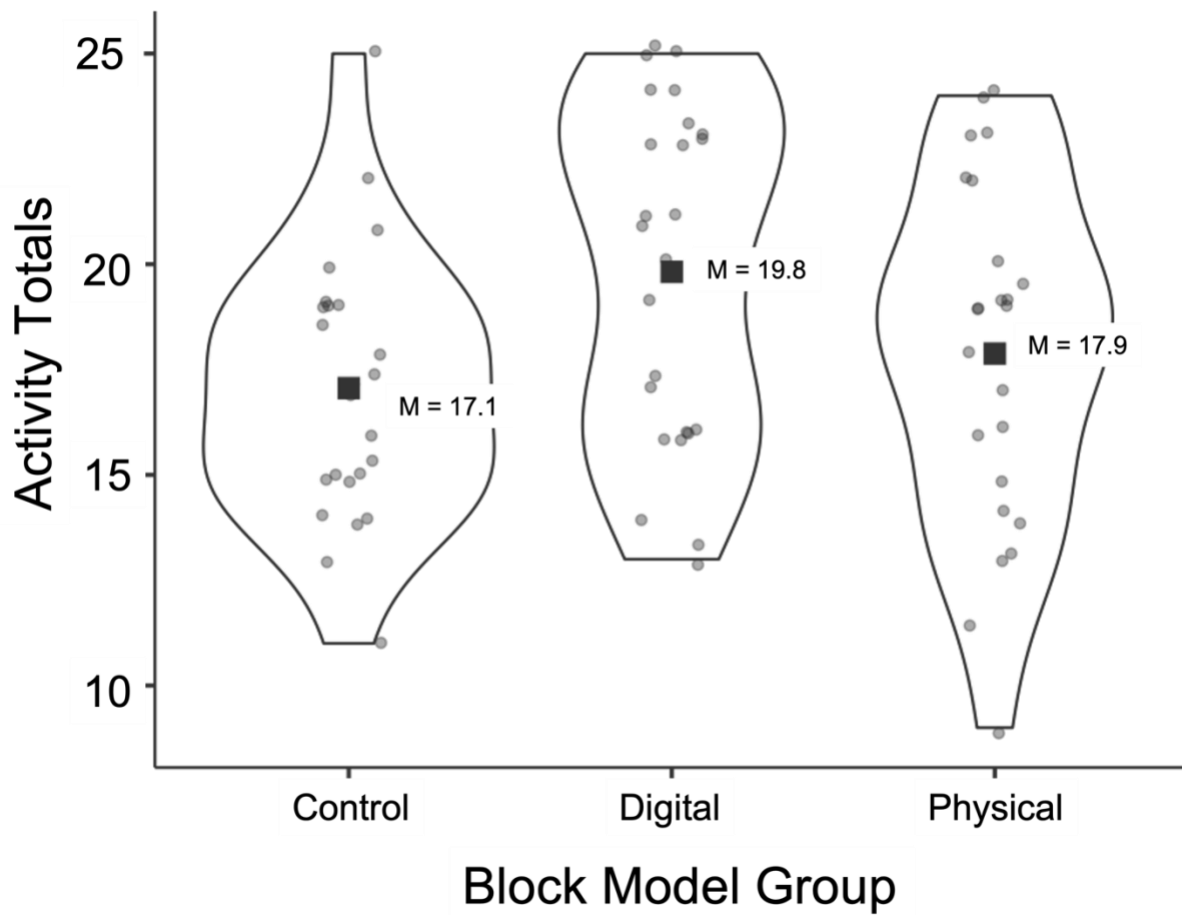


Figure 3. Violin plot displaying student performance on a structural geology activity separated by geologic block model group (control, digital, or physical).

## **Chapter 4**

### **Student Challenges with and Strategies for Solving Visuo-penetrative Structural**

### **Geology Tasks**

## **Abstract**

Spatial visualization skills are used heavily in the geosciences and while there are many studies on measurement and improvement of these skills, there is an opportunity for more investigation into the cognitive processes behind them. Particularly, visual penetration (or penetrative thinking) skills are used in structural geology to understand complex subsurface morphologies based on surficial features. In this study, eye-tracking and retrospective interviews were used to explore the challenges that students encounter when using penetrative thinking to complete a structural geology assessment and the strategies they use to complete the tasks. The primary type of error that students make is penetrative, where they attempt to imagine the inside of an object, compared to nonpenetrative where they focus on surficial stimuli. However, students bring many strategies to solving these problems that can be harnessed for successful problem-solving. This work emphasizes the need for investigating the how of geospatial thinking and collaborative constructivism.

## **Introduction**

Spatial thinking involves the cognitive process of considering objects in space, the relationships between objects, and object transformations. Several studies have explored the role of spatial thinking in Science, Technology, Engineering, and Mathematics (STEM) learning and found that higher spatial reasoning skills are linked to STEM performance and success (Shea et al., 2001; Uttal & Cohen, 2012; Wai et al., 2009). Despite this linkage, spatial reasoning has been neglected in formal educational settings (Frey & Detterman, 2004; Kell & Lubinski, 2013). In spatially-intensive fields, such as the geosciences, robust spatial reasoning is necessary to complete discipline specific tasks (Kastens & Ishikawa, 2006). For example, structural geology tasks require the skill of extrapolating surficial features that extend into the subsurface which cannot be directly observed from the surface. This is an example of within-object relations, a component of spatial thinking (Uttal et al. 2013). If students have less-developed skills, they may have a harder time understanding how to complete geoscience tasks (Ishikawa & Kastens, 2005). Furthermore, lower spatial skills may be a barrier to success for students as they may self-select out of STEM fields (Kell & Lubinski, 2013; Wait et al., 2009) and the geosciences specifically (Rapp et al., 2007; Reynolds et al., 2006; Titus and Horsman, 2009). Thus, infusing explicit spatial learning across the curriculum could broaden the pool of students choosing and succeeding in the geosciences.

### **Characterizing Visuo-penetrative Skills**

There are plenty of examples of researchers measuring spatial thinking skills in the geosciences (e.g., Atit et al., 2015; Gagnier et al., 2107; Ormand et al., 2014; Titus and Horsman, 2009). However, more investigation is needed on how the skills are being activated. Ryker et al. (2018) has set grand challenge priorities for spatial and temporal thinking research which

includes understanding the strategies and processes people use to solve geospatial problems. In a special issue article by Kastens and Ishikawa (2006), the authors delve into the cognitive science literature to elucidate the processes behind solving certain geospatial tasks. They describe a visual processing theory by which there are several steps to taking in information from a 2D object, like perceptions of geometry and depth of the visible surfaces, and creating some 3D mental representation of it. The information gathered by the eye from the visible 2D surfaces of a 3D geologic block model can be useful in describing the features of sides or interiors that are not visible. While this paper does not mention the penetrative thinking skill explicitly, the processes described could still be applicable. The authors recommend future research to explore ways to aid in the 2D to 3D visualization and interpretations, which might include pedagogical tools such as physical or virtual geologic block models.

Penetrative thinking was also investigated and characterized in a study by Kali and Orion (1996) where they analyzed high-school students' performance on structural geology tasks using geologic block models and cross-sections and their beliefs about their skills. Geologic block diagrams (or models) are 3D sections of a bigger geologic structure meant to represent the surface and subsurface morphology including layers of rock, folds, and faults (Kali & Orion, 1996). The researchers developed an instrument consisting of 13 open-ended items subdivided into three categories: drawing cross-sections through block diagrams (four items), completing the other faces of a block diagram given one face (four items), and drawing a third cross section given two others at specified locations on a map (five items). The analysis of student performance and their incorrect answers, particularly, revealed that students are making two primary types of errors: nonpenetrative and penetrative errors. Penetrative errors, errors that were made from attempts to infer the inside of the block diagram, were highly correlated with student

performance where students who made penetrative errors performed better on the instrument than students who made nonpenetrative errors. Nonpenetrative errors are based on the surficial or external information the geologic block provides. Ormand et al. (2014) and McNeal et al. (2021) further describe the penetrative and nonpenetrative errors as either parallel surface, perpendicular surface, or parallelogram error (Figure 1). The parallel surface error is a nonpenetrative error where the block face parallel to the cross-section is inferred to be the cross-sectional surface. The perpendicular surface error is also nonpenetrative and involves interpreting the face perpendicular to the cross-section is the cross-sectional surface. The parallelogram error is penetrative and results in interpreting the cross-sectional surface to have whatever surficial shapes and patterns that exist in the highlighted area. Other errors include straight in (penetrative), wrapping paper (nonpenetrative), and mixed and other which involve some combination of penetrative and nonpenetrative error elements.

Studies that qualitatively probe the process of spatial thinking are few. Furthermore, research has shown that not only are certain skills stronger than others in learners, but that one learner may use a completely different strategy than other learners when solving an identical task (Liben et al., 2012). This study aims to understand and describe the challenges and strategies students present in solving geologic block model problems. The research questions are as follows:

1. What types of visuo-penetrative errors are more common when solving geologic block model problems?
2. What self-reported strategies and visuo-mental processes result in correct answers or common error types?

The spatial thinking skill used in this study is identifying the most likely cross-section of a geologic block diagram from a suite of choices. Geologic block diagrams from the GBCT will be used to assess student-reported challenges and strategies since geologic block diagrams are a common tool for students in geoscience courses.

## **Methods**

### **Theoretical Framework**

We employ a constructivist grounded theory (CGT) approach in the design, implementation, and analysis of this study. CGT is an inductive methodology that emphasizes that reality is socially constructed (Ghezeljeh & Emami, 2009), so the experiences of its members are subjective (Charmaz, 2000; Rieger, 2019). Researchers and participants cocreate the knowledge gained in a CGT study (Charmaz, 2000). CGT is a methodology that can be employed in qualitative inquiry and has rigorous evaluation criteria comparable to that of quantitative inquiry that can be described as validity, reliability, and trustworthiness. Heale and Twycross (2015) define validity as how precisely the concept being studied is measured. In other words, can we say that what we measured is actually measuring the targeted concept, be it a particular spatial thinking skill or an affective measure like engagement. Validity is about truth and certainty, where truth is accurately representing the phenomenon and certainty is being able to support findings with evidence (Schwandt, 2014). A qualitative perspective of validity means achieving validity through methods like triangulation and member-checking (Schwandt, 2014, p. 320). Triangulation, here, means that multiple sources of information or analyses are employed to verify claims made, and member-checking involves checking with participants to verify the researcher's findings (Schwandt, 2014, pp. 193-194). Reliability refers to how consistent or reproducible the results of a study are (Twycross & Shields, 2004). Qualitative inquirers consider

intercoder agreement to be a measure of reliability since thematic and content coding has the potential to be highly interpretive (Creswell and Poth, 2016, p. 264). Intercoder agreement is a measure of how consistently researchers interpret themes and apply them to the narrative data (Kurasaki, 2000). Trustworthiness is more commonly seen in qualitative inquiry. Schwandt (2014) describes trustworthiness as a criterion for judging investigation quality to make it “noteworthy to audiences” (p. 308). The author explains several sub-criteria that include credibility, transferability, dependability, and confirmability. Credibility and transferability ensure that the participants’ views are accurately represented and that the responses and interpretations are generalizable enough to be applied to other similar cases (Schwandt, 2014). Dependability involves heavily documenting the process of inquiry (Schwandt, 2014). Lastly, confirmability means that the data, findings, and further interpretations are “readily discernible” (Schwandt, 2014, p. 309).

### **Participation and Recruitment**

Participants were recruited via email from a variety of disciplines and departments in the undergraduate population of a large-enrollment university in the Southeastern United States. Internal Review Board (IRB) approval for human subjects research was gained prior to commencing the study and students were consented and compensated for participating. Students ( $n = 75$ ) then worked through the 16-question GBCT independently while being eye-tracked to understand where they focus their attention to solve the questions therein. A Tobii TX300 eye-tracker connected to a 23-inch desktop computer monitor collected the data at 300 Hz. Students were calibrated before the data collection began and the calibration results were visually inspected by the researcher to ensure accurate results. Afterward, a random subset of the items on the GBCT were used to conduct a retrospective interview with the students while looking at

replays of their eye-tracked GBCT. Since there was a not a significant difference in pre-post scores (see Chapter 3), the post-GBCT was used for subsequent analyses.

## **Analysis**

Simple central tendency statistics were used to quantify the number of correct and incorrect responses on the GBCT. Incorrect responses were separated according to the type of error that was made and tabulated to understand how common certain errors are among the students in this study. Each answer choice has an error associated with it which was a part of the design of the assessment.

## ***Eye-tracking***

Eye-tracking is used to measure to what and where a person is paying attention by following where the eye is fixating and moving on a screen or an object (Holmqvist et al., 2011). The strengths of using eye tracking are numerous in that they help scientists understand the subconscious cognitive processes that are occurring since our eyes are constantly gathering and making sense of visual information (Tatler et al., 2014). Saccades and fixations are examples of such eye movements. Fixations are points where the eye is still and mental processing takes place (Bojko, 2013). Saccades occur between fixations where a person's eye moves from one fixation point to another (Ramat et al., 2008). Since these eye movements are not conscious, they are considered objective in nature (Tatler et al., 2014). Eye-tracking has been used by geoscientists in many educational settings to evaluate peoples' mental processes for various tasks and skillsets (McNeal et al., 2021). For example, eye-tracking was used to evaluate how graduate and undergraduate students interpreted and understood graphs where graduate students paid more attention to the data within the graph while undergraduate students focused more on graph elements like axes and legends (Atkins & McNeal, 2018). In another climate graph related paper

that employed eye-tracking students engaged more with graphs that incorporated learning science design elements making them more accessible (Courtney & McNeal, 2023). Eye-tracking has also helped researchers (McNeal et al., 2014) evaluate an online curriculum by elucidating how students were engaging with the text and figures so that they can be revised and optimized for use. Another study by Maltese et al. (2014) used eye-tracking headsets with students in a field course to determine how students were engaging with the environment and each other while doing a mapping exercise. Eye-tracking has also been used to investigate student user interactions with online decision support tools focused on climate change and its impacts (Maudlin et al., 2020; Cashwell et al., 2023). These example eye-tracking studies provide insight into student use and processing of earth science material.

Heat maps were generated from participants' eye-tracking results and aggregated according to the answer choice selected on each item. By aggregating, we could qualitatively describe differences in attention to the various elements of the GBCT items that resulted in a correct answer or an error. Five representative items were selected from the following categories based on the type of geological structures represented: dipping beds, upright fold, faulted beds, basin, and plunging folds. Each representative item was selected for use in the heat map analysis. Heat maps were generated with 95% opacity so that underlying stimuli can be seen, 15 pixel radii so that group differences can be more visible, and a relative duration calculation which compares fixation duration to the total fixation durations for the entire time spent on the question slide. Due to the overlapping nature of the areas of interest, interviews were conducted to further elucidate where participants looked.

### ***Interview Data***

Reflexivity via memo writing was used to prevent forcing the researcher's preconceived ideas on the data via the participants (Charmaz 2014; Charmaz 2017). Thematic analysis was employed where emergent themes were identified and analyzed from verbal-response data (Braun and Clarke, 2006; Neuendorf, 2018). Thematic analysis revealed the specific strategy themes that were used to solve the GBCT items. These emergent themes were identified through initial and focused coding. Initial coding involves labeling excerpts with some code, while focused coding takes relevant and frequently occurring codes to categorize larger sections of data (Rieger, 2019). The initial codes were used to label the item that the participant and the researcher were discussing. The number and types of strategies can be quantitatively compared using code co-occurrences, where overlapping codes make it easier to tabulate and summarize large amounts of data (Joffe & Yardley, 2012). This approach helped determine the similar and differing strategies reported between error type. A researcher uninvolved with this project coded 10% of the excerpts from the interview data and deliberated with the primary researcher until a pooled Cohen's kappa score of 0.75 was reached, a substantial agreement score based on Landis and Koch (1977).

### **Results**

The average score on the GBCT was 7.16 out of 16 (or 44.7%) (SD = 4.82). Figure 2 shows a non-normal almost bimodal distribution of scores. There were 664 total incorrect responses that spanned many types of errors (Table 1), but the most counted error was parallelogram (429) followed by parallel face (97). The answer choices with the parallelogram error tended to be either the runner up or the most selected for almost all of the items on the GBCT with the exception of item #13 indicating that this distractor presents a challenge for

students to solving penetrative thinking tasks. McNeal et al. (2021) saw similar results in their investigation of high and low performers.

### **Heat Maps of Representative Items**

Item #1 was used to represent dipping beds and features 5 differently patterned layers of varying thicknesses dipping towards the back left side of the block (Figure 3). Most participants scored correctly (n = 38), but parallelogram error was also highly selected (n = 30) followed by parallel face (n = 6) and perpendicular face (n = 1). There are discernible differences in fixations for each of the answer choices where the correct response has many fixations clustered along the left edge of the bolded cross-section line with minimal fixations elsewhere on the block.

Contrarily the parallelogram heat map shows clusters on the top face and left side of the block along with clustering in the front corner of the block between the bolded cross-section line and the front right face. This could indicate that participants were paying attention to whatever shape they could see in that area and translating that over to the answer choice. The parallel error heat map shows fixation clustering toward the front of the block indicating attention to the front right face of the block. The front right face only displays two shades which can be seen in the answer choice that was selected. The perpendicular error heat map has too few responders to infer with much certainty the strategy that could have been employed.

Item #4 represents the upright fold and features an anticline (or antiform) geologic feature with differently patterned layers of about equal thickness (Figure 4). Again for this block, the participants who answered correctly have fixation clustering very tightly around the left edge of the parallelogram suggesting they are looking at the layers that intersect the cross-section line. The parallelogram and parallel error heat maps show similar fixation distributions. However, the clustering for the parallelogram respondents start at the right edge of the cross-section line and

are more tightly confined to the bounds of the bolded cross-sectional parallelogram. Contrarily, the parallel respondents' fixation clusters mostly start at the front right face of the block and its clusters are not as tightly confined to the cross-sectional bounds. This suggests that the parallelogram respondents are considering the elements of the block that they can see on the surface and interpreting those elements as within the block. Parallel respondents are looking at the block face that is parallel to the cross-section and deciding that the surface is that of the cross-section. Again, with so few respondents in the perpendicular heat map, it is difficult to make sense of the strategy.

Next is item #6, which represents the basin and is the only one of its kind on the GBCT. This item features a cupping of the beds aligned with the center of the block with differentiating symbology of similar thickness (Figure 5). The heat map for the correct respondents again shows fixations around the left edge of the cross-section, though the fixations are more dispersed than for items #1 and #4. However, the fixations are still denser toward the front of the block indicating that they may be getting information from the front and top faces as well. The straight in error heat map seems to be following the same clustering trend, but there are much fewer clusters throughout the front of the block than the correct answer heat map. The respondents likely did not use the other information available on the block aside from the edge to determine the cross section. Straight in is not well defined in the literature, but according to this example it appears that the respondents discovered the beds that would be seen in the cross-section and assumed that the beds would go straight through without any curving. The parallelogram error group have the majority of their fixation clusters within the bounds of the cross section—a pattern seen in the examples before. The parallel error group has clusters more toward the front right face of the block beyond the bounds of the parallelogram potentially indicating that they

were determining which layers could be seen on that face and then selecting the answer choice where those beds are best represented, though the size of the group is comparatively small.

Item #7 represents the plunging fold challenges in the GBCT and features an anticline plunging toward the back left of the block diagram (Figure 6). The beds in the block diagram are differentially labeled and are of about equal thickness. In the correct response group, there appears to be a cluster in every layer across the top and the right side of the block along the cross-section with more intense clustering at the bottom right and top left. Most of the fixations here are within the bounds of the cross-section, but there are scattered clusters toward the front. It appears that the respondents are identifying each layer that needs to be represented as well as the shape they would appear in the cross-section. The parallelogram and straight in error heat maps show heavily concentrated clustering within the bounds of the cross-section, indicating that they didn't use much more information to help solve the item beyond that. The parallelogram group likely saw the curves on the surface of the block and inferred that the inside of the block would look the same. The straight in group likely just identified the layers, but assumed that they would continue straight into the block since they could not identify that the correct shape was an angle. The parallel error group has too few respondents to say with certainty their strategy, but there are isolated pockets of about equal intensity at the front left face and the cross-section boundary.

Lastly, item #15 was selected to represent faulted beds where there are highly folded layers that were subsequently offset by a fault laterally (Figure 7). There is right lateral movement across the fault and the right side is offset higher than the left. The beds are folded in anticline and syncline pairs and have differential patterning and varying thicknesses. Participants who responded correctly focused most of their attention along the center top and top front of the

block likely indicating that they were examining the offsets along the fault and using that to inform their choice. An attention pattern was not entirely apparent from the parallelogram error heat map, but many of the clusters are within the bounds of the cross-section and there is a prominent cluster at the corner of the fault on the block on the right. The respondents were likely determining that the gray corner at the surface must be a part of the cross-section in the same way that it appears at the surface. The parallel error heat map shows intense clustering on the front left face of the block which is parallel to the cross-section. The other error group looks more like perpendicular error based on the answer choice provided since it is showing the waviness of the folds seen on the front right block face. However, the clustering is along the fault from the front left side to about the center of the block running perpendicular to the cross-section's orientation.

Overall, heat maps from correct respondents showed more narrow distributions of fixations compared to other error groups of similar size. The results of these heatmaps complement the scan path data from McNeal et al. (2021).

### **Insights from Interview Data**

To better triangulate the inferences made by the researchers by examining the heat maps, the participants were interviewed to elucidate the strategies employed and determine if they can confirm the inferences. A subset of 30 participants were included in the thematic analysis since saturation was reached quickly. A pooled Cohen's kappa of 0.71 was achieved for inter-coder agreement.

Myriad strategies were reported by participants with varying degrees of success. The strategies are summarized and tabulated based on which aspect of the block models were mentioned in the discussion of their strategy and are delineated by GBCT item (Figure 8). The

numbers in the grid indicate code co-occurrences, or how many times the strategy theme was used for each of the items. Figure 9 shows a codebook of all the codes used, their descriptions, number of code applications, and example excerpts. Most notably participants reported that they used the colors and patterns on the block models, shape of the layers and structures, and corners, borders, and edges most in their strategies. On the other hand, mental measurement, size of the layers, and the interplay between details and the “big picture” were reportedly used less.

Many students reported difficulty with understanding the shapes shown on the block models. As an example, a student who answered incorrectly with a mixed error on item #10 stated,

*“These ones were hard for me. The black lines really threw me off...at first I thought [it would] be a rectangle and then when you look deeper, I was like, it has to be some sort of semicircle. And then with the line incorporated, I just figured it'd be the top left corner cause it was coming from the left side.”*

The student reports trying to see deeper within the block and assumed a difference in the shape that would be shown at the cross-sectional surface.

Students also combined elements in their strategy to solve the block model items. For example, “patterns and colors” and “corners, edges, and borders” had a high code co-occurrence meaning many students paid attention to those elements in their strategy. One student remarked on item #2 that they answered correctly,

*“I'm just looking...seeing like the side first, [because] that'll help me get like a, a reference image about what I'm supposed to be looking at. So for (answer choice) C, I'm not noticing any like irregularities that go like from here out into it so I can kind of eliminate that*

*one. And then with D I noticed that this portion has a little bit of really dark gray, so automatically I can kind of tell that D's gonna be that correct answer."*

The student was able to rule out certain distractors by identifying which beds would be present in the cross-section at the edge of the cross-section shape.

Another combination with high code co-occurrence was "block faces" and "shape". This student commented that "*looking at the different layers and then seeing what shapes they made at the intersection of the line across*" was their overall strategy when asked by the researcher.

The student was able to determine the internal shape by deciding which block face was needed to establish the shape that would continue through the block to the location of the cross-section.

## **Discussion**

The benefit of employing constructivism in the data collection and analysis processes allows researchers to use the existing mental frameworks that students bring to educational spaces. These results show that students are already thinking about ways to solve these spatial challenges, and many are doing so successfully. Researchers and practitioners can use the elements that already stand out to students, like recognizing fault displacement or identifying and matching patterns, to better build the skills needed to teach structural geology and other geoscience topics.

### **Students' Mental Visualizations**

The results show a bimodal distribution of scores similar to that seen in Kali and Orion (1996), suggesting that the same mental barriers in "seeing" the inside of an object that existed then still persist today carrying over from high school to college students. From the eye-tracking and interview data, students who are consistently making nonpenetrative errors are unable to use

all of the contextual information on the block to visual the inside. Instead, they are stopping at one face of the block and assuming it is identical to the cross-sectional surface. This was apparent from the eye-tracking heat maps the discreet areas that students paid attention to. Students who had more correct or penetrative errors answers discussed in their interviews that they were viewing multiple and more specific areas of the block employing more of a process. Students making other nonpenetrative errors seemed to look at the shapes and colors of one face to make their answer choice selection and reported shorter, less process-based descriptions of their strategies. These students making nonpenetrative errors need special attention and diverse approaches to ensure that these errors are corrected in hopes to also improve performance on other geospatial tasks (Ishikawa & Kastens, 2005).

## **Recommendations**

Based on the findings from this study, several co-created strategies have been developed for solving the major types of geologic block model problems like the ones seen in the GBCT. Eye-tracking data reported here have elucidated differences in students' approaches and process of working through a penetrative thinking challenge, while interview data have further distinguished the key elements that can be used in spatial training. The primary strategy discovered is knowing where on the block diagram to go to look for what information and using what we can see first before trying to imagine within. Some tips generated from the results are as follows:

1. Identify the orientation of the cross-section relative to the block. This will help to eliminate the perpendicular error because the cross-section will not be cutting in that direction.

2. Identify all of the layers that touch the cross-section including their relative thicknesses and patterns. This will rule out any answer choices that do not have all of the necessary layers included.
3. Identify the block face that is visible and parallel to the cross-section. This will give a sense of the shape the layers will likely make in the cross-section. This should help mitigate the parallelogram error by focusing attention away from the stimuli that appear to be within the cross-section but are just on the surface. It will also prevent the straight in error by showing exactly what shape should be in the cross-section instead of assuming perpendicular lines.
4. Locate any faults and intrusions as they may affect the position and thickness of the layers in the cross-section.

A visualization of these strategies can be found in Figure 9.

### **Limitations**

With any study employing qualitative methods, the interpretations have some bias to the knowledge, skills, and expertise of the researcher. Additionally, uncertainty is introduced in any form of self-reporting. Students were given a multiple-choice spatial thinking assessment meaning they only had a small number of items to select from, and therefore the full range of strategy possibilities could not be explored in this study.

### **Conclusion**

Penetrative thinking skills are fundamental in structural geology and other geoscience disciplines. While there have been many studies exploring ways to measure and improve the skill and the theoretical frameworks, research is limited on students' mental processes solving penetrative thinking tasks. In this study, we use eye-tracking technology and interview data in

collaboration with students to explore those strategies and processes. The findings from this study firstly emphasize the need for student collaboration and constructivism in the research process. Secondly, they provide a basis for improving geospatial training for low-performing students making penetrative and nonpenetrative errors in structural geology problems.

### **Acknowledgements**

The authors would like to thank the Geocognition and Geoscience Education Lab members of Auburn University. Also, we appreciate Dr. Carol Ormand for providing the GBCT assessment for use in this study.

### **Funding**

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1414475. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

### **Disclosure Statement**

The authors report there are no competing interests to declare.

## References

- Atit, K., Gagnier, K., & Shipley, T. F. (2015). Student gestures aid penetrative thinking. *Journal of Geoscience Education*, 63(1), 66-72.
- Atkins, R. M., & McNeal, K. S. (2018). Exploring Differences Among Student Populations During Climate Graph Reading Tasks: An Eye Tracking Study. *Journal of Astronomy & Earth Sciences Education (JAESE)*, 5(2), 85-114.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research In Psychology*, 3(2), 77-101.
- Charmaz, K. (2000). Grounded Theory: Objectivist and Constructivist Methods. In Denzin, N. & Lincoln, Y.(Eds.) *Handbook of Qualitative Research. Thousand Oaks, California: Sage*, 509-535.
- Charmaz, K. (2014). *Constructing grounded theory*. Sage.
- Charmaz, K. (2017). Special invited paper: Continuities, contradictions, and critical inquiry in grounded theory. *International Journal of Qualitative Methods*, 16(1), 1609406917719350.
- Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, 15(6), 373-378.
- Ghezjeljeh, T. N., & Emami, A. (2009). Grounded theory: methodology and philosophical perspective. *Nurse Researcher*, 17(1).
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-based Nursing*, 18(3), 66-67.
- Ishikawa, T., & Kastens, K. A. (2005). Why some students have trouble with maps and other

- spatial representations. *Journal of Geoscience Education*, 53(2), 184-197.
- Joffe, H., & Yardley, L. (2012). Content and thematic analysis. In D. F. Marks & L. Yardley (Authors), *Research Methods For Clinical And Health Psychology* (pp. 56-66). Los Angeles, CA: Sage.
- Kali, Y., & Orion, N. (1996). Spatial abilities of high-school students in the perception of geologic structures. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 33(4), 369-391.
- Kastens, K. A., & Ishikawa, T. (2006). Spatial thinking in the geosciences and cognitive sciences: A cross-disciplinary look at the intersection of the two fields. *Special Papers- Geological Society of America*, 413, 53.
- Kell, H. J., & Lubinski, D. (2013). Spatial ability: A neglected talent in educational and occupational settings. *Roeper Review*, 35(4), 219-230.
- Kurasaki, K. S. (2000). Intercoder reliability for validating conclusions drawn from open-ended interview data. *Field Methods*, 12(3), 179-194.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
- Liben, L. S., & Titus, S. J. (2012). The importance of spatial thinking for geoscience education: Insights from the crossroads of geoscience and cognitive science. *Geological Society of America Special Papers*, 486, 51-70.
- Lobeck, A. K. (1924). *Block Diagrams and Other Graphic Methods Used In Geology And Geography*. J. Wiley.
- Maltese, A. V., Balliet, R. N., & Riggs, E. M. (2013). Through their eyes: Tracking the gaze of students in a geology field course. *Journal of Geoscience Education*, 61(1), 81-88.

- McNeal, K. S., Atkins, R., & Johnson, E. T. (2021). Visualizing Student Navigation of Geologic Block Diagrams. In *Applying Bio-Measurements Methodologies in Science Education Research* (pp. 295-308). Cham: Springer International Publishing.
- McNeal, K. S., Libarkin, J. C., Ledley, T. S., Bardar, E., Haddad, N., Ellins, K., & Dutta, S. (2014). The role of research in online curriculum development: The case of EarthLabs climate change and Earth system modules. *Journal of Geoscience Education*, 62(4), 560-577.
- Neuendorf, K. A. (2018). Content analysis and thematic analysis. In *Advanced research methods for applied psychology* (pp. 211-223). Routledge.
- Ormand, C. J., Manduca, C., Shipley, T. F., Tikoff, B., Harwood, C. L., Atit, K., & Boone, A. P. (2014). Evaluating geoscience students' spatial thinking skills in a multi-institutional classroom study. *Journal of Geoscience Education*, 62(1), 146-154.
- Rapp, D. N., Culpepper, S. A., Kirkby, K., & Morin, P. (2007). Fostering students' comprehension of topographic maps. *Journal of Geoscience Education*, 55(1), 5-16.
- Reynolds, S. J., Piburn, M. D., Leedy, D. E., McAuliffe, C. M., Birk, J. P., & Johnson, J. K. (2006). The Hidden Earth—Interactive, computer-based modules for geoscience learning. *Geological Society of America Special Papers*, 413, 157-170.
- Rieger, K. L. (2019). Discriminating among grounded theory approaches. *Nursing Inquiry*, 26(1), e12261.
- Ryker, K., Jaeger, A. J., Brande, S., Guereque, M., Libarkin, J., & Shipley, T. F. (2018). "Research on Cognitive Domain in Geoscience Learning: Temporal and Spatial Reasoning". In St. John, K (Ed.) (2018). *Community Framework for Geoscience*

*Education Research*. National Association of Geoscience Teachers.

[https://doi.org/10.25885/ger\\_framework/7](https://doi.org/10.25885/ger_framework/7)

Schwandt, T. A. (2014). *The Sage dictionary of qualitative inquiry*. Sage publications.

Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology*, 93(3), 604.

Twycross, A., & Shields, L. (2004). Validity and reliability-What's it all about? Part 2 Reliability in quantitative studies. *Paediatric Nursing*, 16(10), 36.

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why, and how?. In *Psychology of Learning and Motivation* (Vol. 57, pp. 147-181). Academic Press.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101(4), 817.

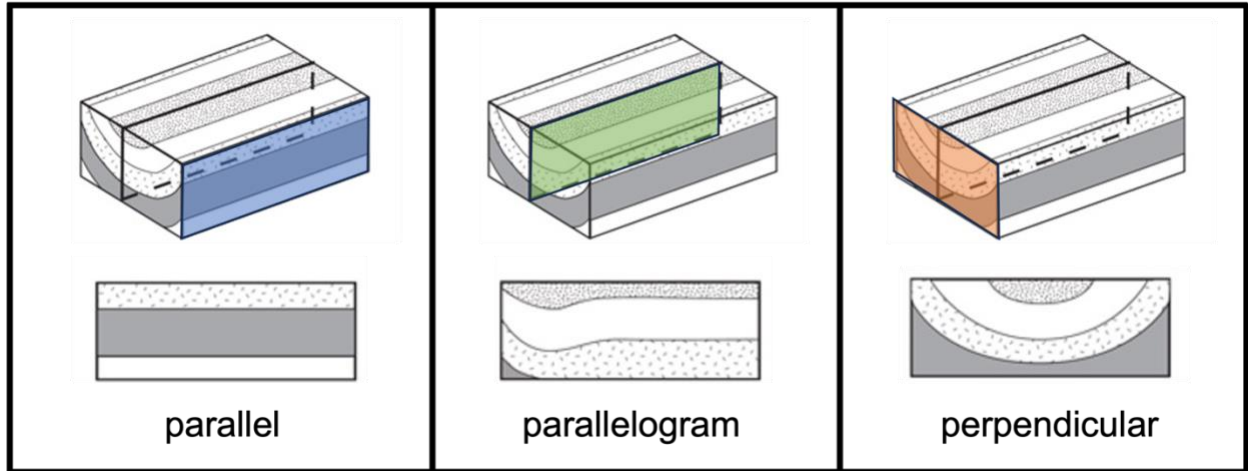


Figure 1. Three examples of common errors that are made when solving for the cross-section in the block model. Parallel and perpendicular errors are nonpenetrative errors, while parallelogram errors are penetrative.

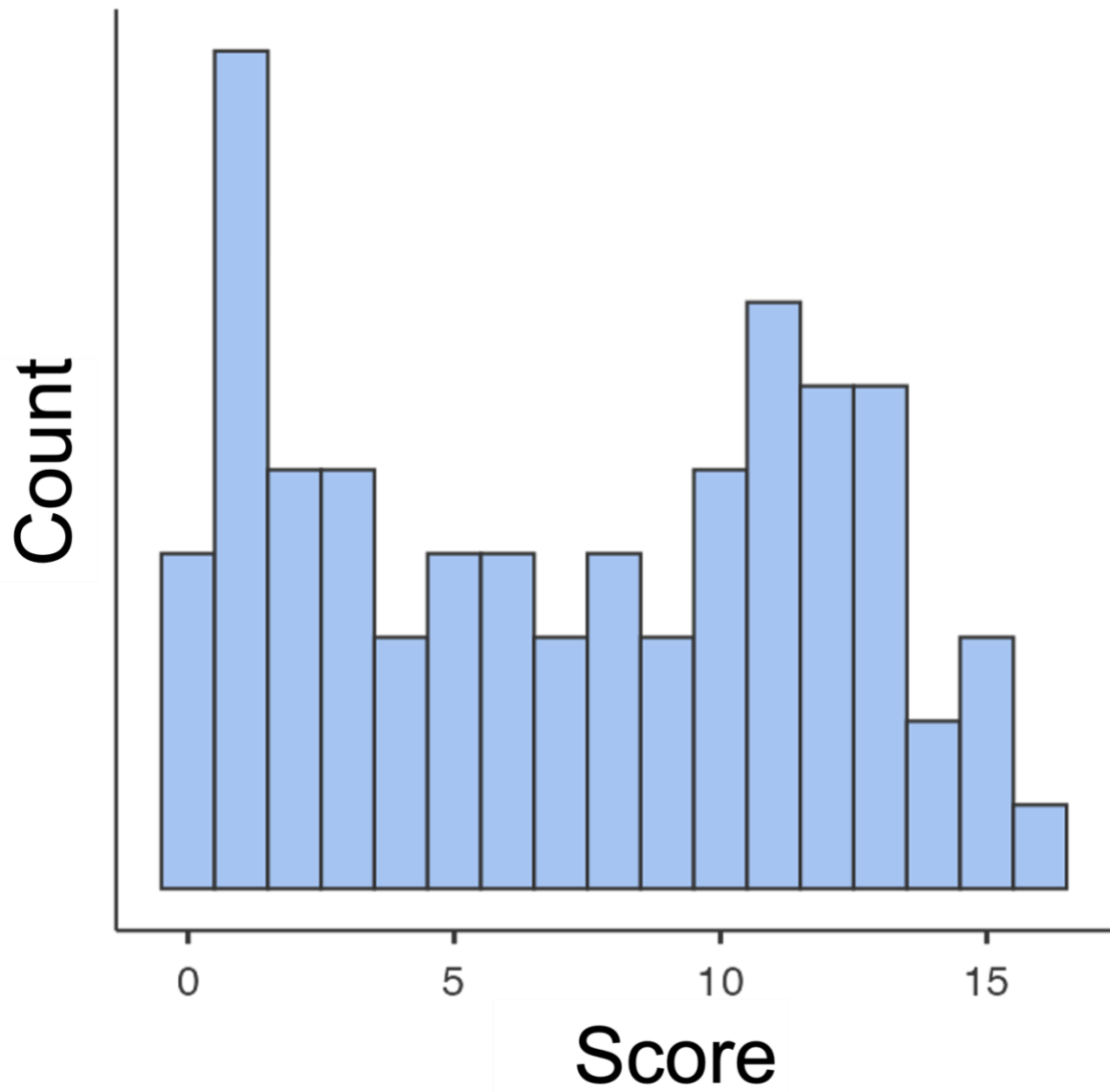


Figure 2. Histogram showing the distribution of scores on the GBCT.

Table 1.

*Correct and incorrect responses separated by item and by error type. COR = correct count, %COR = percent correct, PG = parallelogram error, P = parallel error, PP = perpendicular error, SI = straight in error, WR = wrapping error, M = mixed error, OTH = other error, INC = incorrect count, %INC = percent incorrect, and Avg = average.*

Item	COR	%COR	Error Type							INC	%INC
			PG	P	PP	SI	WR	M	OTH		
1	38	50.7	30	6	1					37	49.3
2	43	57.3	25	3					4	32	42.7
3	43	57.3	21	2					9	32	42.7
4	44	58.7	22	7	2					31	41.3
5	44	58.7	15	3	13					31	41.3
6	33	44.0	24	4		14				42	56.0
7	33	44.0	34	3		5				42	56.0
8	19	25.3	53	3						56	74.7
9	22	29.3	39	6				8		53	70.7
10	40	53.3	28	3					4	35	46.7
11	23	30.7	28	3		21				52	69.3
12	20	26.7	39	10		6				55	73.3
13	13	17.3	16	15					31	62	82.7
14	25	33.3	26	16		8				50	66.7
15	49	65.3	12	8					6	26	34.7
16	47	62.7	17	5		6				28	37.3
	536	44.7	429	97	16	60	8	35	19	664	55.3
	Total	Avg								Total	Avg

## Item #1. Dipping Beds

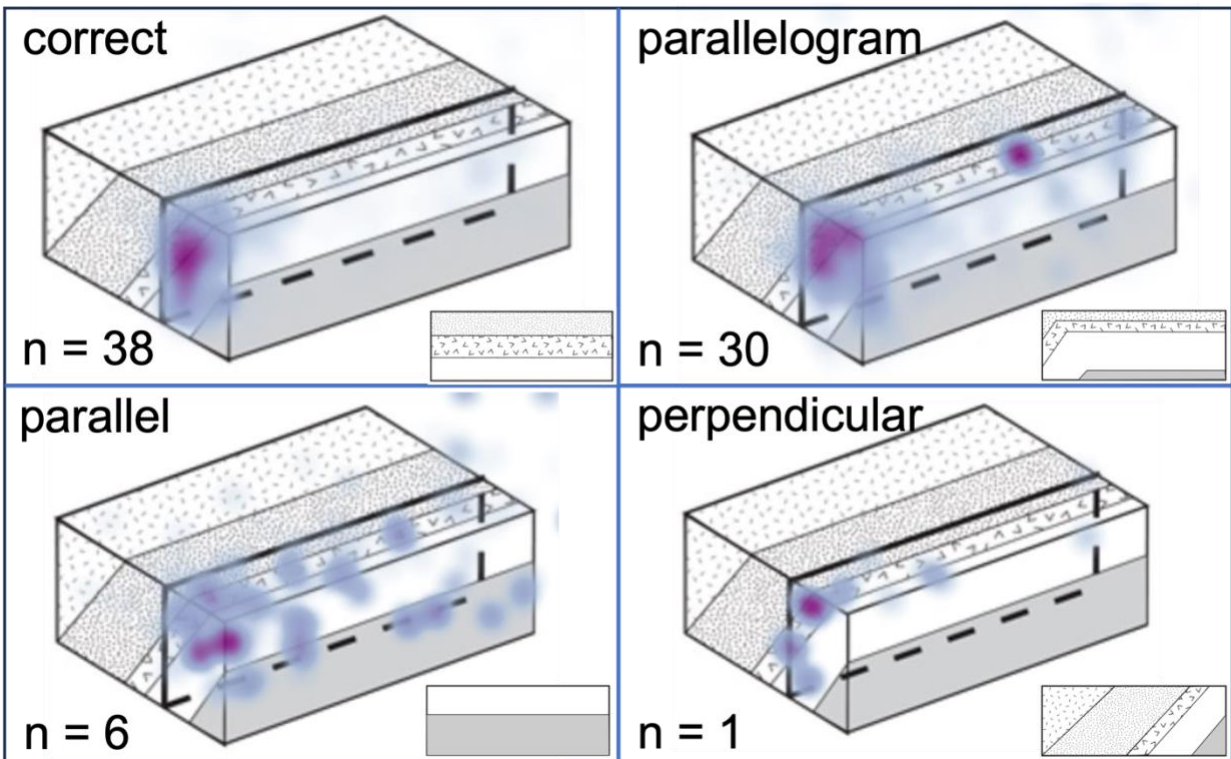


Figure 3. Heat maps of the item #1 from the GBCT to represent dipping beds. Heat maps are aggregated by participants who selected the answer choice (inset in the bottom right of the square) that corresponds to the correct or incorrect options. N = the number of participants who selected the answer choice.

## Item #4. Upright Fold

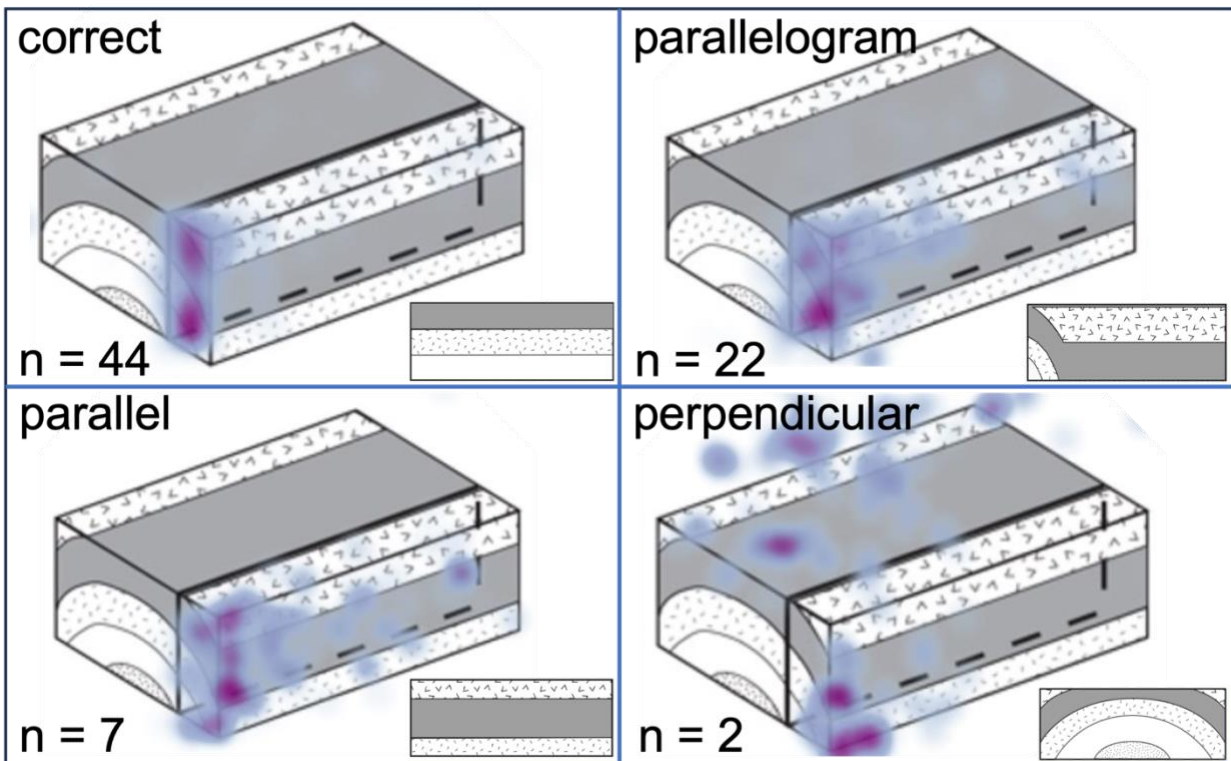


Figure 4. Heat maps of the item #4 from the GBCT to represent upright folds. Heat maps are aggregated by participants who selected the answer choice (inset in the bottom right of the square) that corresponds to the correct or incorrect options. N = the number of participants who selected the answer choice.

## Item #6. Basin

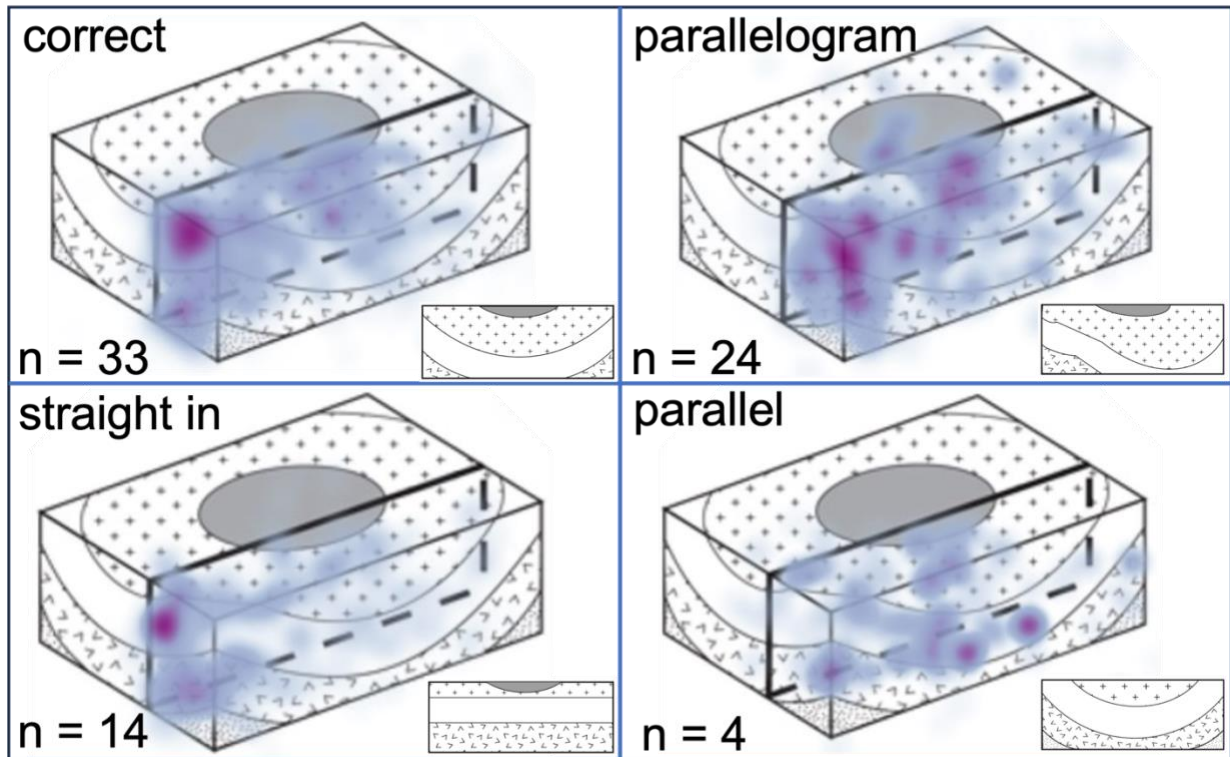


Figure 5. Heat maps of the item #6 from the GBCT to represent basins. Heat maps are aggregated by participants who selected the answer choice (inset in the bottom right of the square) that corresponds to the correct or incorrect options. N = the number of participants who selected the answer choice.

## Item #7. Plunging Fold

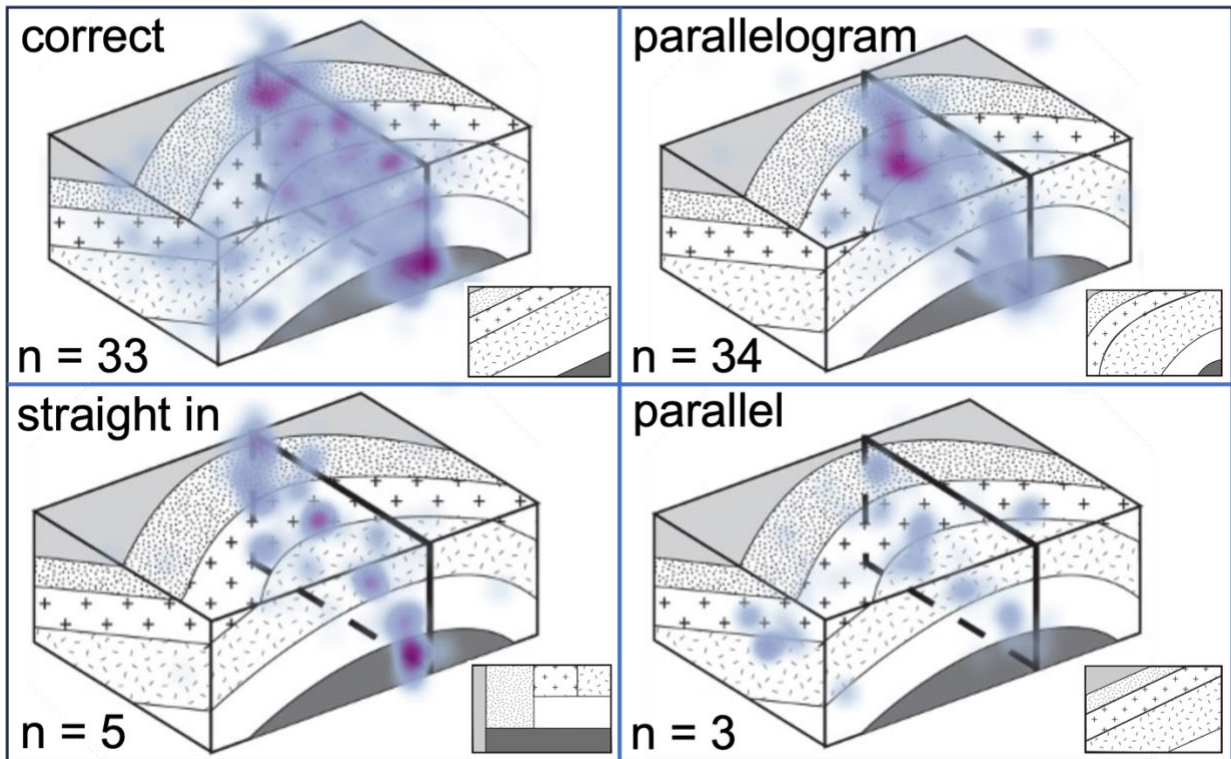


Figure 6. Heat maps of the item #7 from the GBCT to represent plunging folds. Heat maps are aggregated by participants who selected the answer choice (inset in the bottom right of the square) that corresponds to the correct or incorrect options. N = the number of participants who selected the answer choice.

## Item #15. Faulted Beds (with Fold)

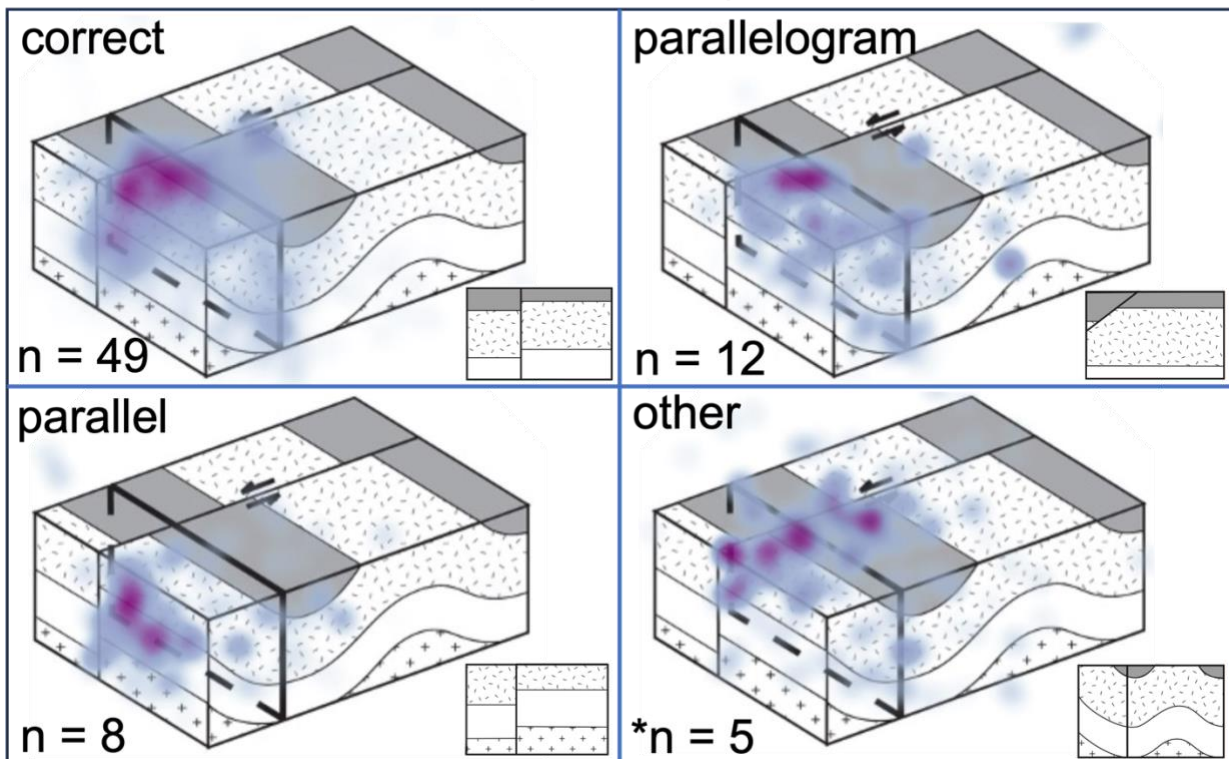
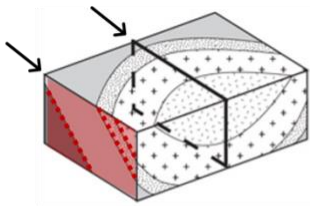


Figure 7. Heat maps of the item #15 from the GBCT to represent faulting beds. Heat maps are aggregated by participants who selected the answer choice (inset in the bottom right of the square) that corresponds to the correct or incorrect options. N = the number of participants who selected the answer choice. \*One of the respondents' data diminished the intensity of the others likely from error during data collection, so their eye-tracking data were removed from this heat

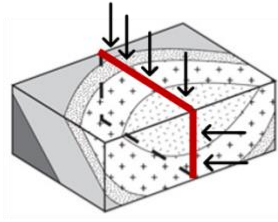
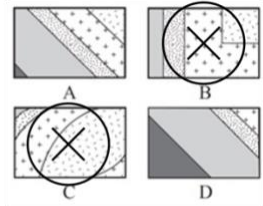
map.

	big picture>details	block faces	complex areas	corners, borders, and edges	cross-cutting relationships	details>big picture	extrapolating through block	layers	movement	patterns and colors	perspective-taking	shape	size	Totals
Q overall		2	1	2		1	1	1	2	3		4		17
Q unlabeled				1			1		1	2	1			6
Q01		4	1	2	1		1	4		8	1	5	1	28
Q02				3				2		4		3		12
Q03				1										1
Q04		1				1	1		1	1	1	2		8
Q05		1		1	1			2	1	6	2	2		16
Q06								1		1		2		4
Q07		2		1				1	1	4	1	5		16
Q08	1	1	1	2	1		1	3	1	5		7	1	26
Q09				2			1			2	1	2	1	9
Q10		4	2	1	2	1	1			1		4		17
Q11								1		1	1	1	1	5
Q12			1	3				1		3	1	2		11
Q13	1			1	1	1	1			2		5		12
Q14		3		1				1		3		2		11
Q15	1	1		1			2	1	3	4		4	1	19
Q16									2	2	1			5
big picture>details				1					1	1		1	1	8
block faces			2	1	3	1	2		2	7		12		49
complex areas		2		2					1	1	1	3		16
corners, borders, and edges	1	1	2					4	1	13	3	5	1	53
cross-cutting relationships		3				1	1	2		3	1	3		20
details>big picture		1			1		1					2		9
extrapolating through block		2			1	1			1	6		3		24
layers				4	2				1	14	3	9	1	52
movement	1	2	1	1			1	1		5	1	4	1	30
patterns and colors	1	7	1	13	3		6	14	5		6	23	2	133
perspective-taking			1	3	1			3	1	6		5	1	31
shape	1	12	3	5	3	2	3	9	4	23	5		2	122
size	1			1				1	1	2	1	2		14
Totals	8	49	16	53	20	9	24	52	30	133	31	122	14	

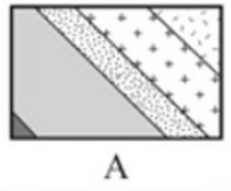
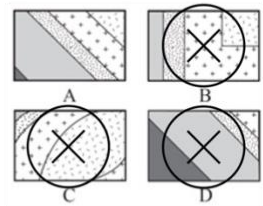
Figure 8. Code co-occurrences grid displaying the key elements of the block model and themes students discuss as a part of their strategy for solving the GBCT items. Red colors indicate more code co-occurrences, green is medium, and blue is few.



1. Identify that the cross-section is parallel to the front left face of the parallelogram. Notice that the layers are angled from the top left to the bottom right. This will be the orientation of the layers in the cross-section. This eliminates answer choices B and C.



2. Identify all of the patterns along the solid visible portion of the cross-section. Note their relative thicknesses and their order along the top and right side of the block diagram. A is the only choice that has all of the patterned layers along the cross-section line.



Answer choice A is the correct answer because it has the layers angled diagonally in the correct orientation and has all of the patterns that intersect the cross-section line.

Figure 9. Example solution for a geologic block model challenge.

Table 2.

*Codebook showing the codes created, the number of code applications, the description of the code and an example excerpt for each.*

<b>Code</b>	<b>App Count</b>	<b>Description</b>	<b>Excerpt Example</b>
big picture > details	2	ignoring the details to understand the overall conformation	"I'm seeing like what is the cross-section cut through, touching that line, but then like, just looking at the full picture to see what's included in that box."
block faces	19	focusing on specific sides or faces	"So I was thinking we're like looking at it from this side of it. So I'm trying to think which one I chose. I probably chose this one...just looking at this side."
complex areas	5	considering areas of highest complexity	"So find the constant thing... then from there, ID the things that may look different or...more complex, complex areas."
corners, borders, and edges	20	describing stimuli along corners, borders, and edges	"I mean, you're clearly gonna have some of this on the edge. Okay. So, and it looks, and to me it looks like it's going away. Definitely dark gray on the edge, and it's probably going away."
cross-cutting relationships	6	describing features that truncate or overlay others	"I think I started by looking at the intrusion part just because obviously it's folded. It just seemed easier to go from there."
details>big picture	4	focusing on specific features rather than the overall conformation	"So I won't look at the answers, but I'll stare at one spot and be like, how would that look?"
extrapolating through block	10	discussing the elements that continue or discontinue into the surface	"And then just referencing off of where this line begins and ends, I can kind of anticipate the depths of the dark gray."

layers	18	mentioning layered material, beds, strata, or layers	"I mean, Yeah, I guess I wanted to make sure I had all the layers, right?"
movement	11	inferring movement or displacement (e.g., faults)	"So with this one, because there is a fault, there's a change in the material when it slides. So I can't rely on the old [tactic] that I used, like going on top, because that material's gonna be completely different."
patterns and colors	52	identifying patterns and colors	"Like, I know I'm gonna see this section...and sort of...discerning each section by, you know, like, this one's just gray. This one has like grittiness. So like, I know I'm gonna see these sections, whichever ones are touching the line."
perspective-taking	10	describing viewpoints, angles, and ways to "see" the different block elements	"I just figured that seeing from right there would be better at figuring out... I guess just looking, like, you can just see it at all the different angles."
shape	48	referencing shapes or conformations of the block elements	"Um, looking at the different layers and then seeing what shapes they made at the intersection of the line across."
size	4	referring to size, thickness, or scale of the block elements	"So the bottom or the top of the triangle would be like, the smallest amount that you would see, like for one of the answers, and it would slowly get bigger and like wider as you go out."

## **Chapter 5**

# **Geospatial Thinking: Conclusions and Reflections on Spatial Thinking in the Geosciences**

Spatial thinking has been a topic of interest for researchers and practitioners for years and efforts to measure, explore usage, and support it in the geosciences has required collaboration from cognitive scientists, geoscience education researchers, and instructors from various academic levels and geoscience disciplines. The projects in this dissertation contribute to and expand these efforts by explicitly addressing the measurement, use, and support priority areas.

### **Measurement**

In this study, a survey was developed to understand how certain beliefs about and experiences with spatial thinking compared between STEM and non-STEM college graduates. The findings showed that experience, confidence, and measured spatial skills are not aligned for STEM and non-STEM graduates where STEM graduates reported having higher confidence and more experience, but performed worse than non-STEM graduates. Despite this, there was still a strong major to career consistency suggesting that spatial skills were not the main determinant of success nor persistence in STEM after college. This should inform more research into other diverse measures that more holistically assess career pathways. Future directions of this research include expanding the SRCES survey questions to include more types of experiences than the few listed. Also, the spatial thinking skill assessment used, the SRI, only measures three skills: mental rotation, spatial orientation, and spatial visualization.

### **Support**

This study explored the potential of physical and digital block models to improve spatial visualization skills, particularly penetrative thinking, and affect performance on a structural geology activity over using flat topographic maps. The need for this project was emphasized by two geoscience practitioners who were curious about how different pedagogical tools could affect student outcomes. We expected to see better performance on both the spatial assessment

and the assessment by using one of the block models compared to not using one. While the physical block group showed the most improvement from pre to post assessment, the result was not significant. However, there was a significant difference in activity performance between the digital block model and control groups. Both the physical and digital block model groups outperformed the control group. This study has major implications for practitioners' access to digital visual aids and resources since physical visual aids can be inaccessible and impractical for many and traditional aids show lower efficacy in student performance.

## Use

Continuing in the vein of penetrative thinking, this study involved investigating specific student strategies for and challenges to solving items on the GBCT. Eye-tracking and retrospective interviews were used to explore students' mental processes. Constructivism was a central theoretical framework with the assumption that the best group to help create pedagogy is the very group we will use these pedagogies on. Findings include elucidating specific portions of the block models that students paid attention to as well as descriptions of the points and how they were used in solving the questions. A general guide to interpreting cross-sections from block models informed by the students' data was created for practitioners to use in their instructions. This study emphasizes that there are multiple ways of solving penetrative thinking problems, so it is important for students' perspectives to be consulted in the research process to co-create pedagogy.

My hope is that researchers continue to collaborate with practitioners and students to co-create and test tools, interventions, and processes to improve the state of geoscience education.

## Supplementary Materials A. Spatial Reasoning Confidence and Experience Survey

### (SRCES) Questions

#### Confidence Questions

4 = Strongly agree, 3 = Somewhat agree, 2 = Somewhat disagree, 1 = Strongly disagree

**C1** <- I feel confident that my answers on this assessment are correct.

**C2** <- I feel confident that I did better than most people on this survey.

**C3** <- I thought this survey was easy.

#### Demographics

**Age** <- Participant age

**Gender** <- Participant gender

**Ethnicity** <- Participant ethnicity

#### Education and Career

**Education, EC1** <- What was your highest degree attained

**Major, EC2\_1** <- If you have a college degree, select your college major. If you do not see yours represented, select the closest field.

**Major2, EC2\_2** <- College major subfield

**Career, EC3\_1** <- Select the area that best represents your current job sector. If you do not see yours represented, select the closest sector.

**Career2, EC3\_2** <- Current job subfield

### Spatial Experience

5 = Always, 4 = Frequently, 3 = Sometimes, 2 = Rarely, 1 = Never

**SE1** <- How often do you rely on spatial reasoning in your job?

**SE2** <- How often do you rely on spatial reasoning in your daily life?

**SE3** <- How often do you rely on automatic instructions from a GPS service (i.e. Google Maps) to navigate where you want to go? \*dropped

**SE4** <- How often do you rely on a paper map to navigate where you want to go?

**SE5** <- How often do you use navigation tools like a compass to navigate where you want to go (not computerized GPS service)?

**SE6** <- How often did you play with building blocks or Legos as a child? \*dropped

**SE7** <- How often did you play action video games as a child?

**SE8** <- How often do you play action video games as an adult?

5 = At least once a week, 4 = Once a month, 3 = Once or twice a year, 2 = Not very often, 1 = Never

**SE9** <- How often do you engage in outdoor activities that require some amount of navigation?

### Spatial Confidence

4 = Strongly agree, 3 = Somewhat agree, 2 = Somewhat disagree, 1 = Strongly disagree

**SC1** <- I do well when given a task that requires me to imagine and manipulate a spatial environment

**SC2** <- I do better than others at tasks that require me to imagine and manipulate a spatial environment.

**SC3** <- I do well finding locations on a map when given compass directions.

**SC4** <- I do well finding locations when a computerized GPS service is giving me directions. \*dropped

**SC5** <- I do well locating myself on a map. \*dropped

**SC6** <- I do well rotating objects in my head.

# Supplementary Materials B. Geologic Block Model Activity (No Block Model)

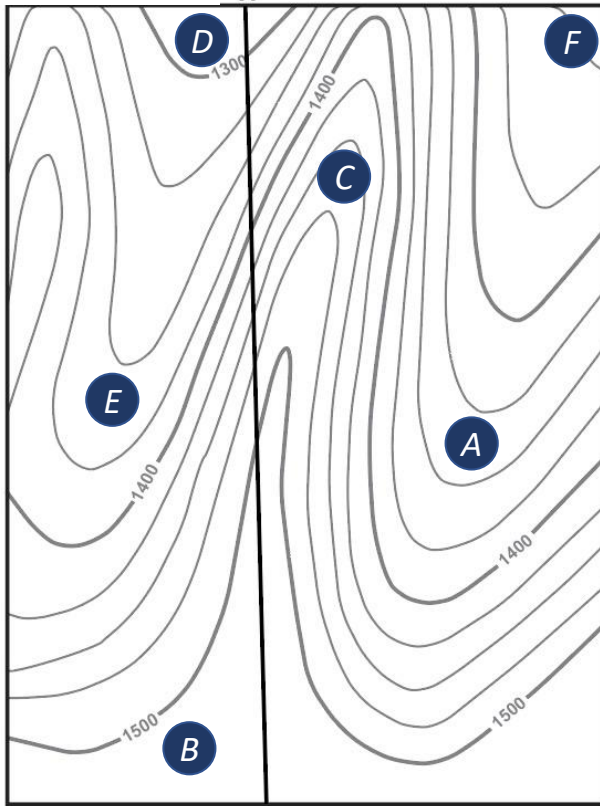
## EXERCISE ONE



**Learning Goal:** Identify contour lines and sketch topographic profiles from elevation data.

Examine the topographic map below and locate the contour interval. This number means that there are 25 meters of elevation between each contour line (in gray).

### Block 88D



100 m  
Contour Interval = 25 m

### Questions



1. Locate where points A, B, C, D, E are on the topographic map.

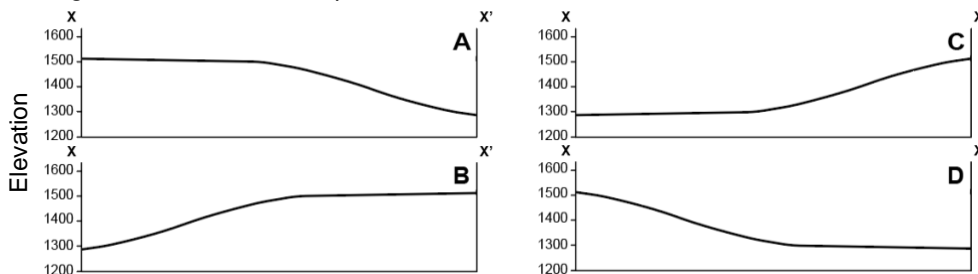
Which point is located at the **highest** elevation? \_\_\_\_\_

Which point is located at the **lowest** elevation? \_\_\_\_\_

How do you know where the highest and lowest points are located?

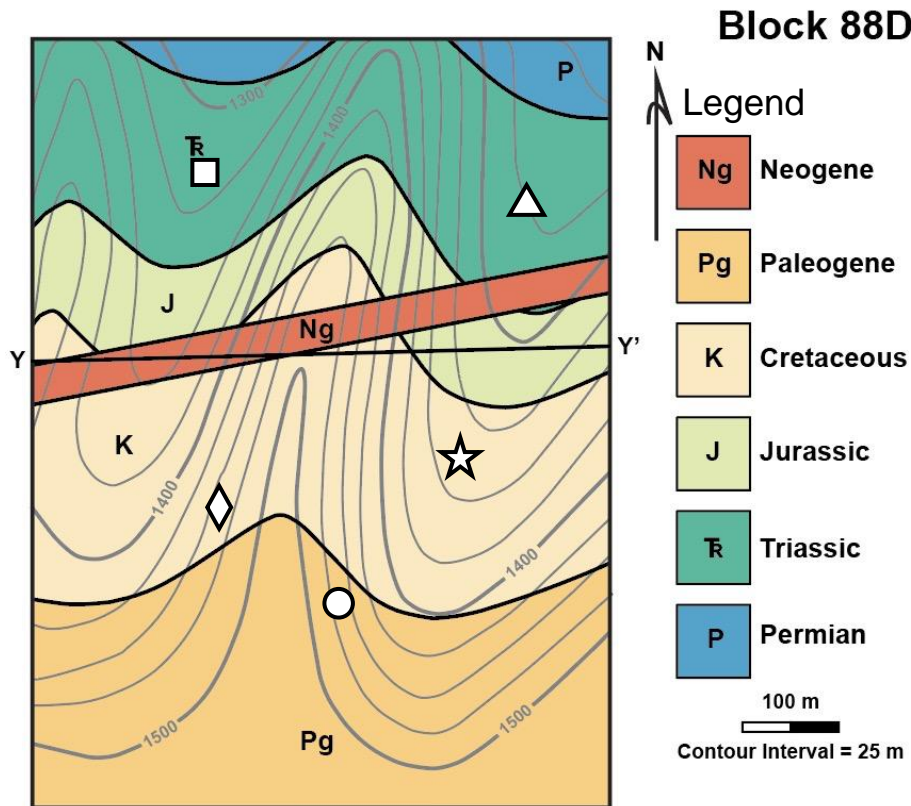
2. Assuming there is no vegetation that could block your view and it is nighttime, if there was a campfire located at point E, at which of the other points would an observer be able to see the light from the fire? How do you know?

3. Trace your finger along the path of X-X' on the map. Which one of the following is the correct topographic profile along the line from X at the top to X' at the bottom?



## EXERCISE ONE

The map below shows the same topography as the previous, but we have added a little more information. The legend now shows several rock units of various ages. In the legend, the units are ordered from youngest at the top to oldest at the bottom. Notice that there are shapes on this map to represent specific locations. Locate the square, triangle, star, diamond, and circle.



4. Assuming that there is no vegetation to block your view and it is nighttime, if there was a campfire located at the star, at which of the other positions (square, triangle, diamond, circle) would an observer be able to see the light from that fire? How do you know?

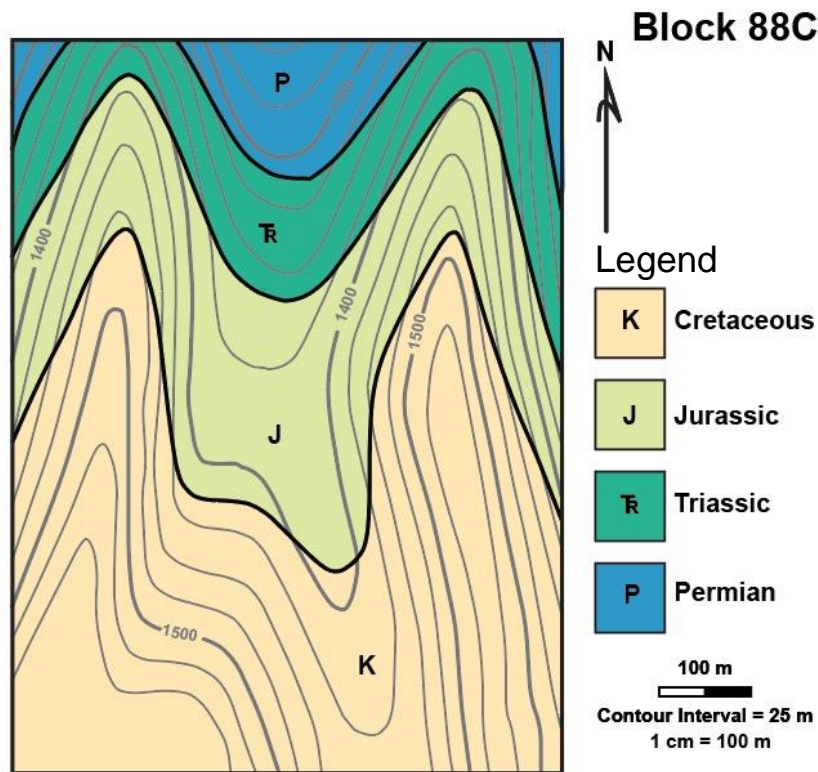


5. Trace your finger along the path of Y-Y'. Sketch a topographic profile along the line from Y on the left to Y' on the right. Use the chart to the left to sketch the topographic profile.

## EXERCISE TWO

**Learning Goal:** Identify and differentiate between the contacts and faults on geologic maps.

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines).

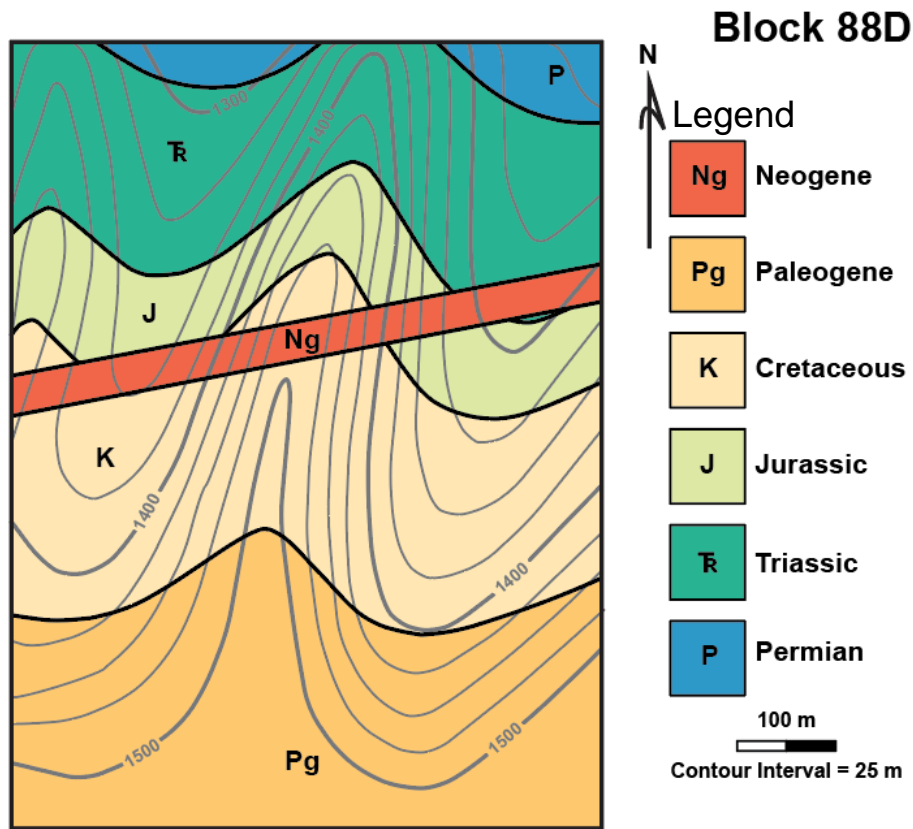


### Questions

1. Describe the appearances of the contour lines and the contacts. How would you describe their relationship?
2. By examining the contacts between the Jurassic and Triassic units, how steeply do these **units** dip? Shallow or steep? How do you know?
3. In what direction do these **units** dip? Use the North arrow provided with the map to help answer this question.

## EXERCISE TWO

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). Identify the feature of Neogene age that cuts across these beds.



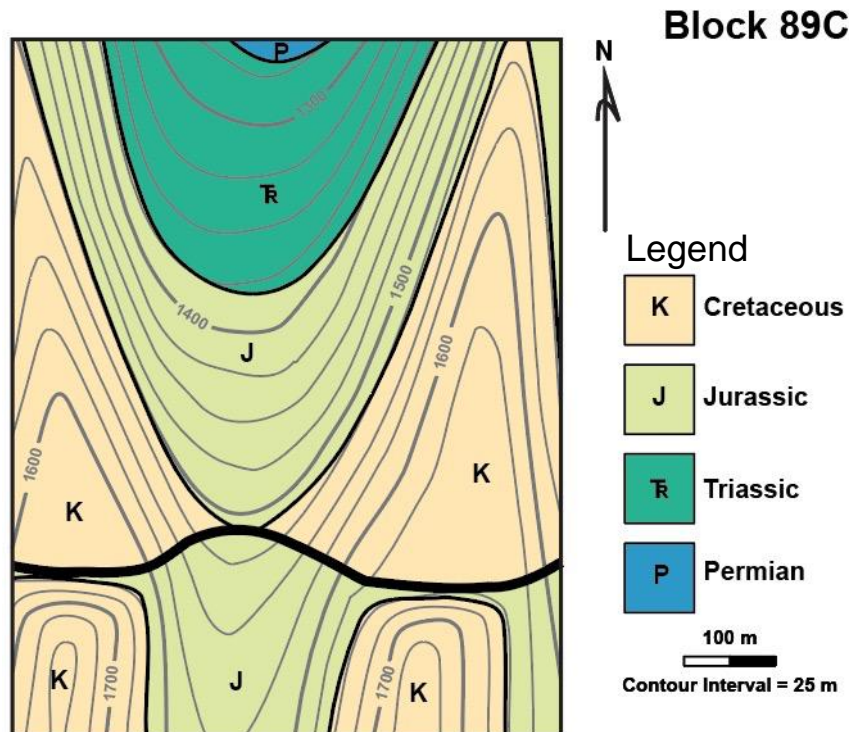
4. By examining the contacts between the Jurassic and Triassic units, how steeply do these **units** dip? Shallow or steep? How do you know?

5. There is a contact between the younger Neogene unit and the rest of the units on the map. Is this **contact** steeply or shallowly dipping?

### EXERCISE THREE

**Learning Goal:** Estimate the dip of the geologic units and structures and determine their direction of dip from geologic map data.

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). Now, we are adding a new thicker black line to indicate a fault. Label the fault on the map below.

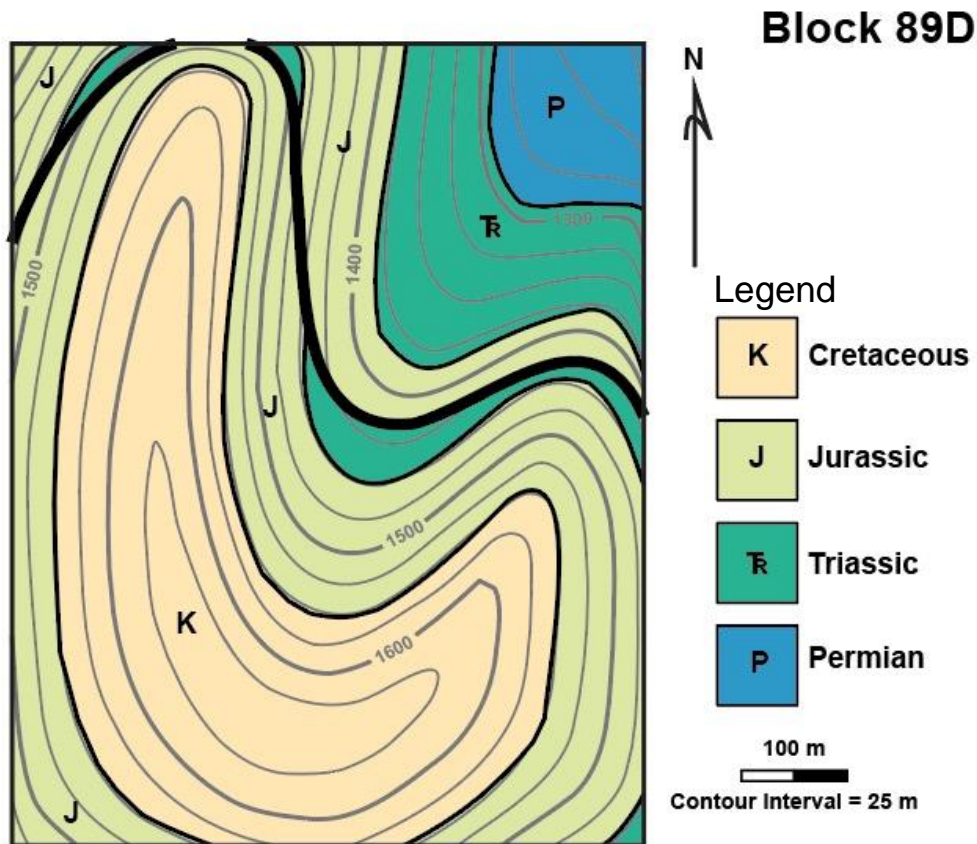


#### Questions

1. By examining the contacts between the Cretaceous and Jurassic units, how steeply do these **units** dip? Shallow or steep? How do you know?
2. Examine the fault and explain whether it is dipping shallowly or steeply.
3. What direction does the **fault** dip in? Use the North arrow on the map to help you answer.
4. What kind of fault is this? Normal, reverse/thrust, or strike-slip? How do you know?

### EXERCISE THREE

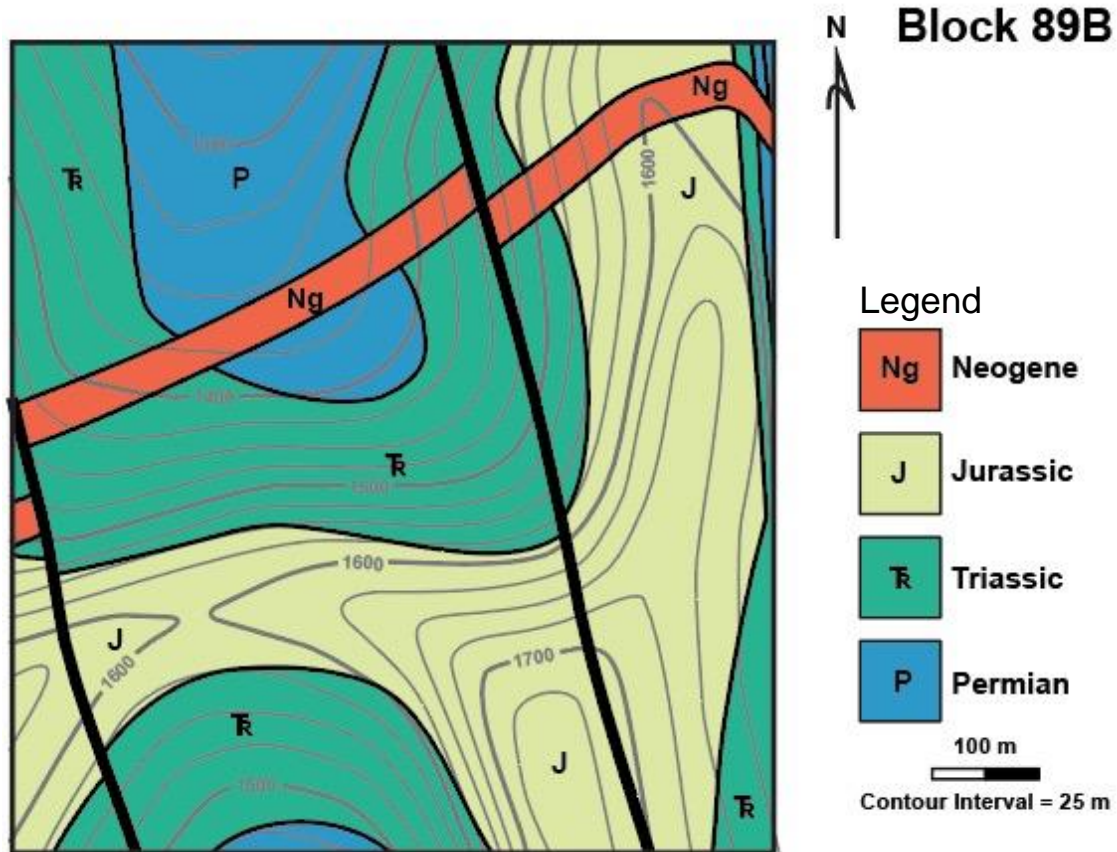
Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). The thicker black line indicates a fault. Label the fault on the map below.



5. By examining the contacts between the Cretaceous and Jurassic units, how steeply do these **units** dip? Shallow or steep? How do you know?
  
6. Examine the fault and explain whether it is dipping shallowly or steeply.
  
7. What direction does the **fault** dip in? Use the North arrow on the map to help you answer.
  
8. What kind of fault is this? Normal, reverse/thrust, or strike-slip? How do you know?

### EXERCISE THREE

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). The thicker black line indicates a fault. Label the faults on the map below.



9. By examining the contacts between the Triassic and Jurassic units, how steeply do these **units** dip? Shallow or steep? How do you know?

10. Examine the fault and explain whether it is dipping shallowly or steeply.

11. What direction does the **fault** dip in? Use the North arrow on the map to help you answer.

12. What kind of fault is this? Normal, reverse/thrust, or strike-slip? How do you know?

# Supplementary Materials C. Geologic Block Model Activity (Digital Block Model)

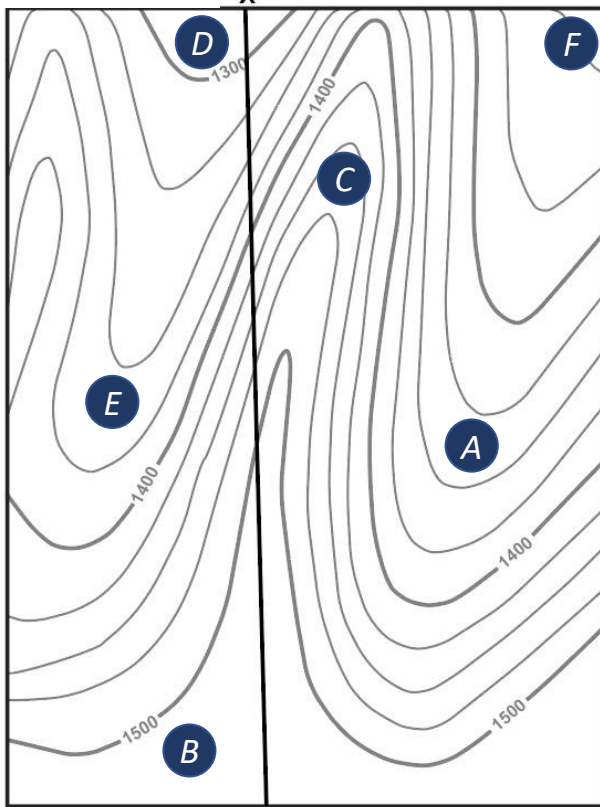
## EXERCISE ONE



**Learning Goal:** Identify contour lines and sketch topographic profiles from elevation data.

Examine the topographic map below and locate the contour interval. This number means that there are 25 meters of elevation between each contour line (in gray). First, align the digital block with the surface shown below.

### Block 88D



100 m  
Contour Interval = 25 m

### Questions

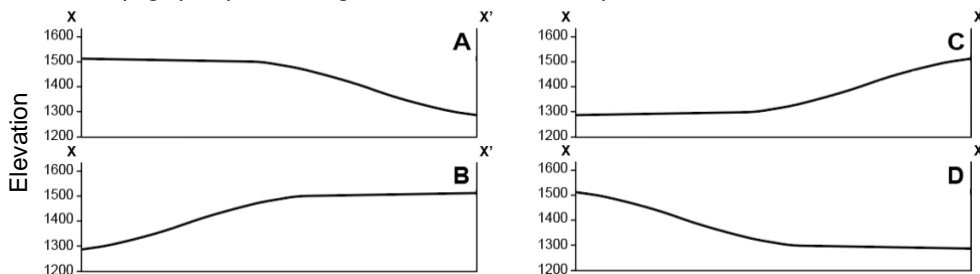


1. Locate where points A, B, C, D, E are on the topographic map and where they would approximately be on the digital block model. Which point is located at the **highest** elevation? \_\_\_\_  
Which point is located at the **lowest** elevation? \_\_\_\_

How do you know where the highest and lowest points are located?

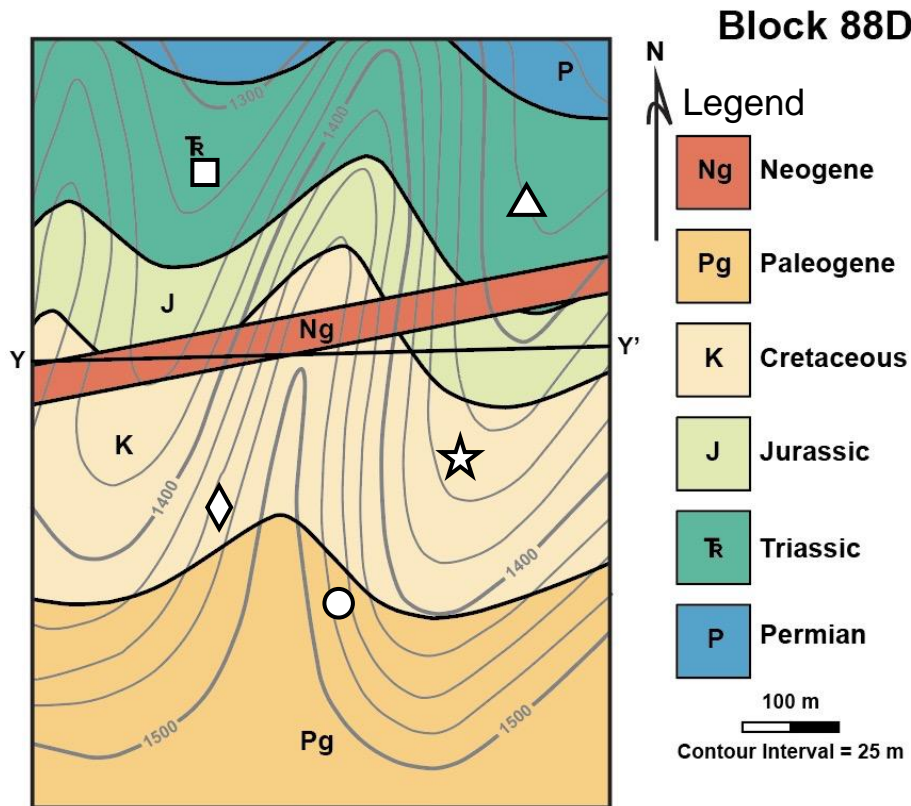
2. Assuming there is no vegetation that could block your view and it is nighttime, if there was a campfire located at point E, at which of the other points would an observer be able to see the light from the fire? How do you know?

3. Trace your cursor along the path of X-X' on the map and the digital block model. Which one of the following is the correct topographic profile along the line from X at the top to X' at the bottom?



## EXERCISE ONE

The map below shows the same topography as the previous, but we have added a little more information. The legend now shows several rock units of various ages. In the legend, the units are ordered from youngest at the top to oldest at the bottom. Notice that there are shapes on this map to represent specific locations. Locate the square, triangle, star, diamond, and circle on the 2D map and its approximate location on the digital model.



4. Assuming that there is no vegetation to block your view and it is nighttime, if there was a campfire located at the star, at which of the other positions (square, triangle, diamond, circle) would an observer be able to see the light from that fire? How do you know?

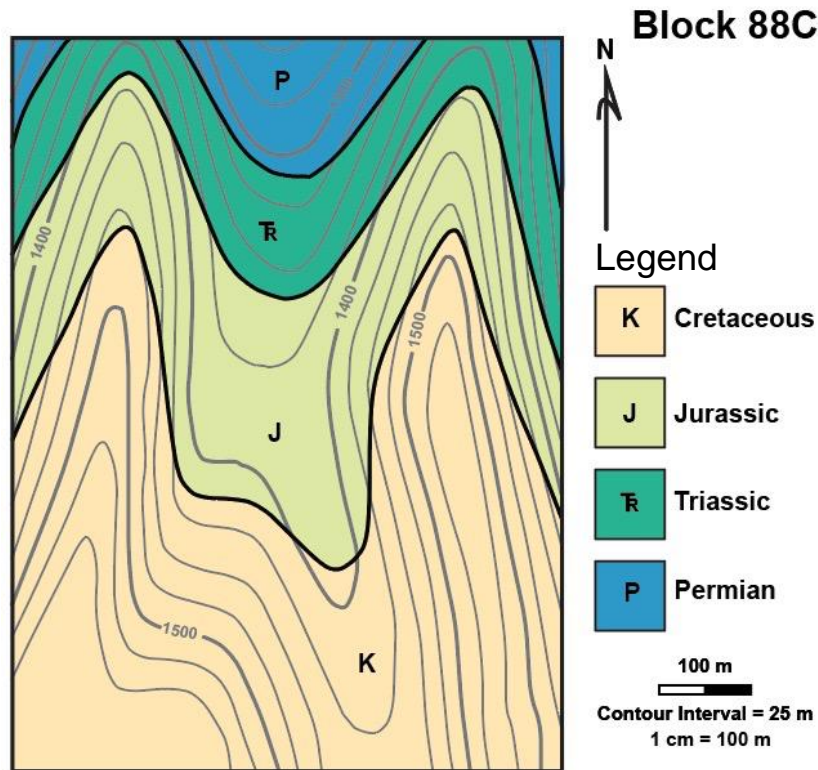


5. Trace your cursor along the path of Y-Y'. Sketch a topographic profile along the line from Y on the left to Y' on the right. Use the chart to the left to sketch the topographic profile.

## EXERCISE TWO

**Learning Goal:** Identify and differentiate between the contacts and faults on geologic maps.

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). Use your cursor to trace the contacts on the digital model.

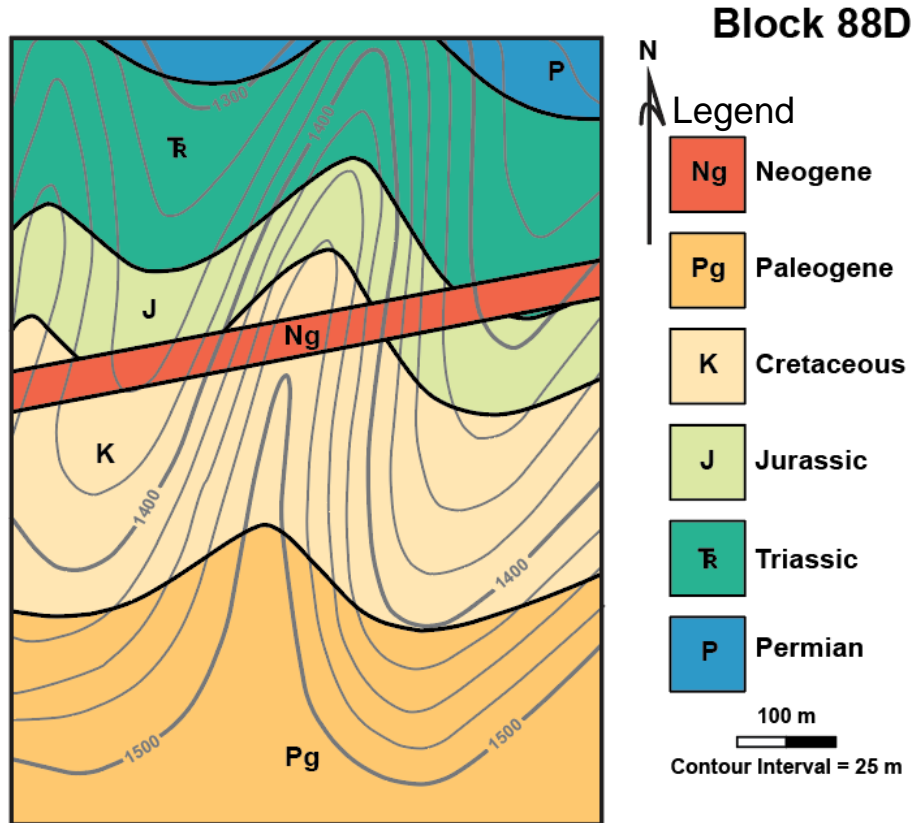


### Questions

1. Describe the appearances of the contour lines and the contacts. How would you describe their relationship?
2. By examining the contacts between the Jurassic and Triassic units, how steeply do these **units** dip? Shallow or steep? How do you know?
3. In what direction do these **units** dip? Use the North arrow provided with the map to help answer this question.

## EXERCISE TWO

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). Identify the feature of Neogene age that cuts across these beds. Trace this feature with your cursor:



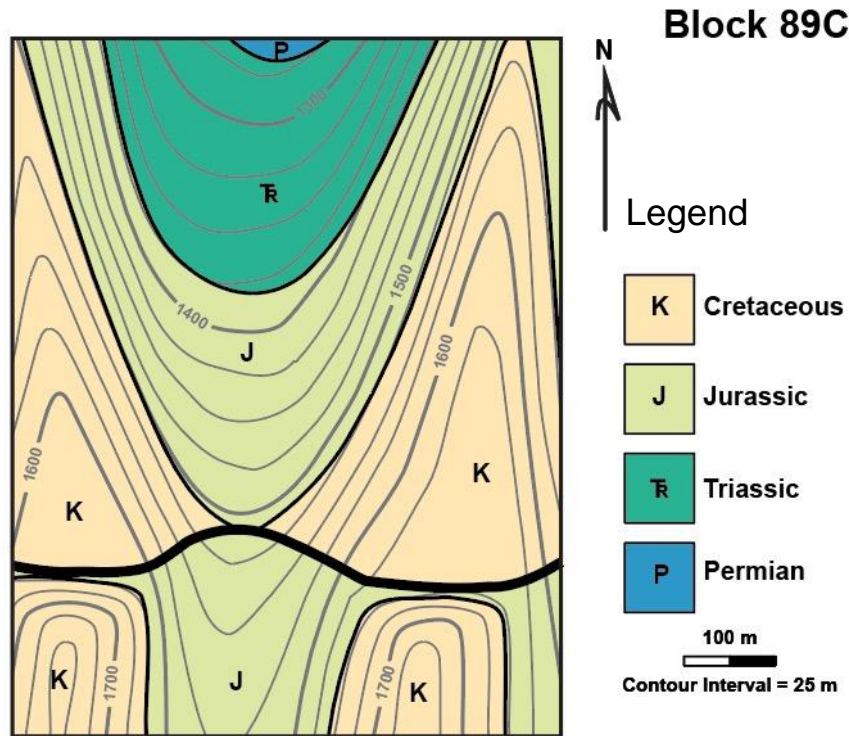
4. By examining the contacts between the Jurassic and Triassic units, how steeply do these **units** dip? Shallow or steep? How do you know?

5. There is a contact between the younger Neogene unit and the rest of the units on the map. Is this **contact** steeply or shallowly dipping?

### EXERCISE THREE

**Learning Goal:** Estimate the dip of the geologic units and structures and determine their direction of dip from geologic map data.

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). Now, we are adding a new thicker black line to indicate a fault. Label the fault on the map below. Use your cursor to trace the **fault** on the digital model. Identify the difference between the appearance of contacts and faults.

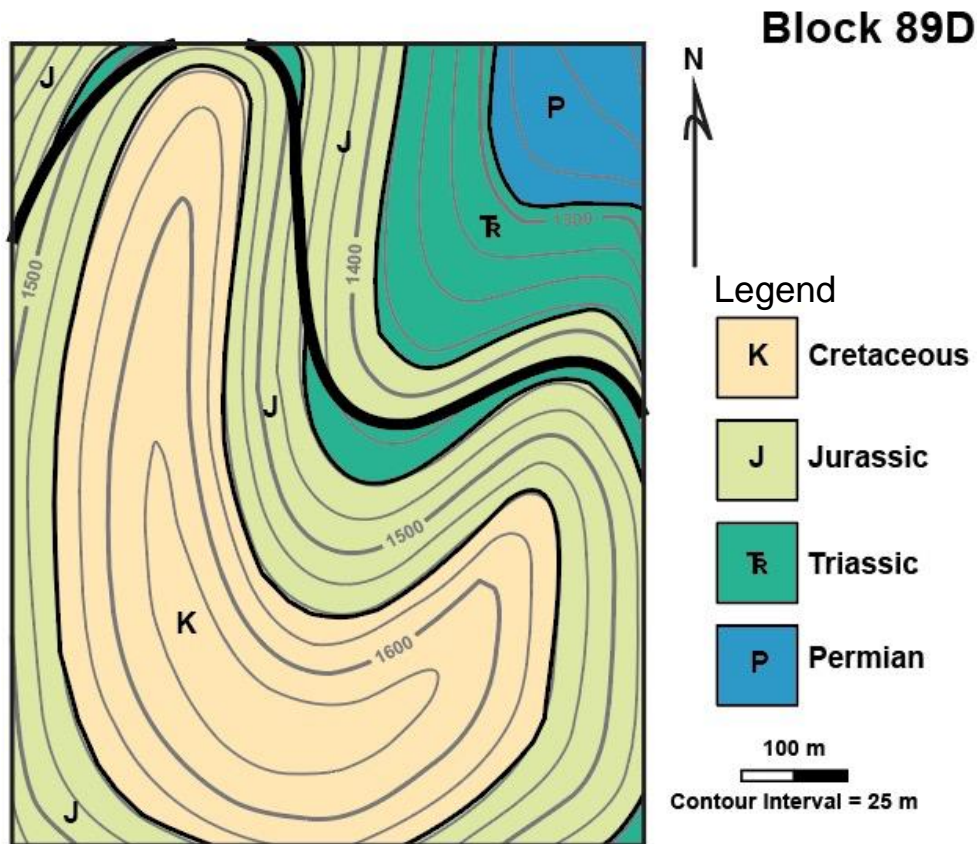


#### Questions

1. By examining the contacts between the Cretaceous and Jurassic units, how steeply do these **units** dip? Shallow or steep? How do you know?
2. Examine the fault and explain whether it is dipping shallowly or steeply.
3. What direction does the **fault** dip in? Use the North arrow on the map to help you answer.
4. What kind of fault is this? Normal, reverse/thrust, or strike-slip? How do you know?

### EXERCISE THREE

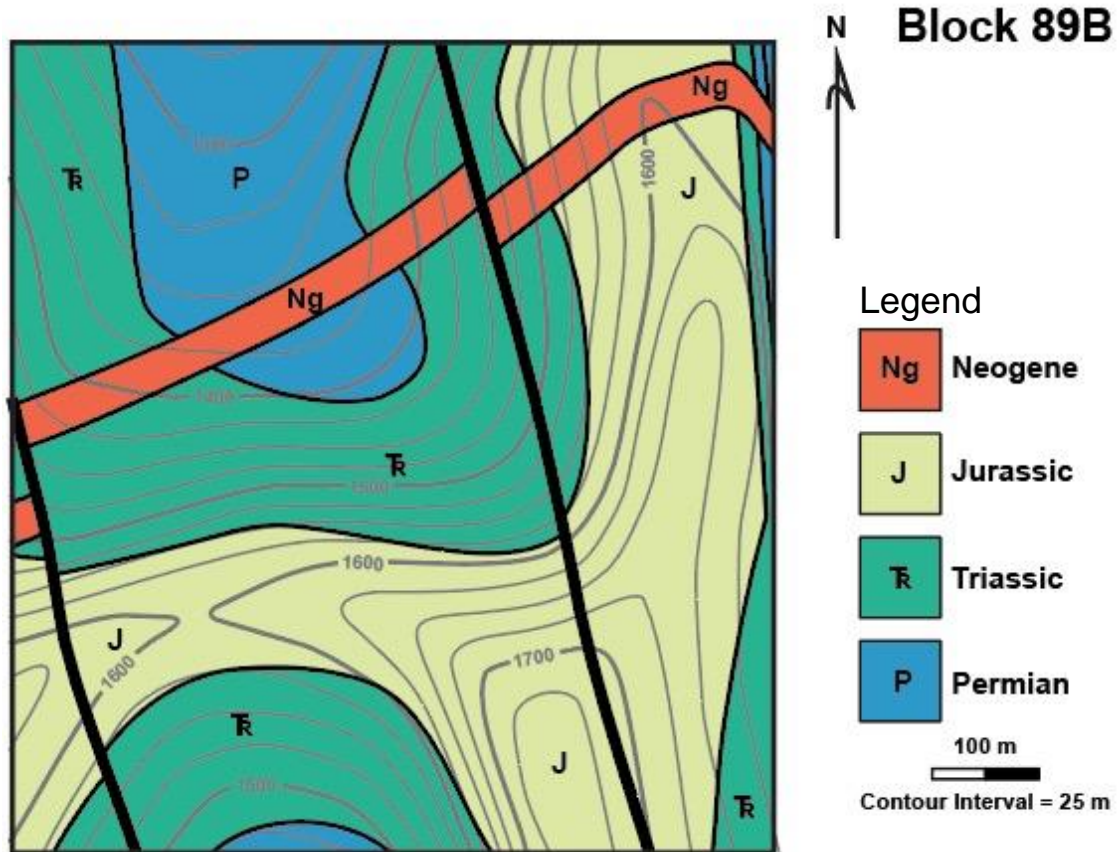
Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). The thicker black line indicates a fault. Label the fault on the map below.



5. By examining the contacts between the Cretaceous and Jurassic units, how steeply do these **units** dip? Shallow or steep? How do you know?
  
6. Examine the fault and explain whether it is dipping shallowly or steeply.
  
7. What direction does the **fault** dip in? Use the North arrow on the map to help you answer.
  
8. What kind of fault is this? Normal, reverse/thrust, or strike-slip? How do you know?

### EXERCISE THREE

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). The thicker black line indicates a fault. Label the faults on the map below.



9. By examining the contacts between the Triassic and Jurassic units, how steeply do these **units** dip? Shallow or steep? How do you know?

10. Examine the fault and explain whether it is dipping shallowly or steeply.

11. What direction does the **fault** dip in? Use the North arrow on the map to help you answer.

12. What kind of fault is this? Normal, reverse/thrust, or strike-slip? How do you know?

# Supplementary Materials D. Geologic Block Model Activity (Physical Block Model)

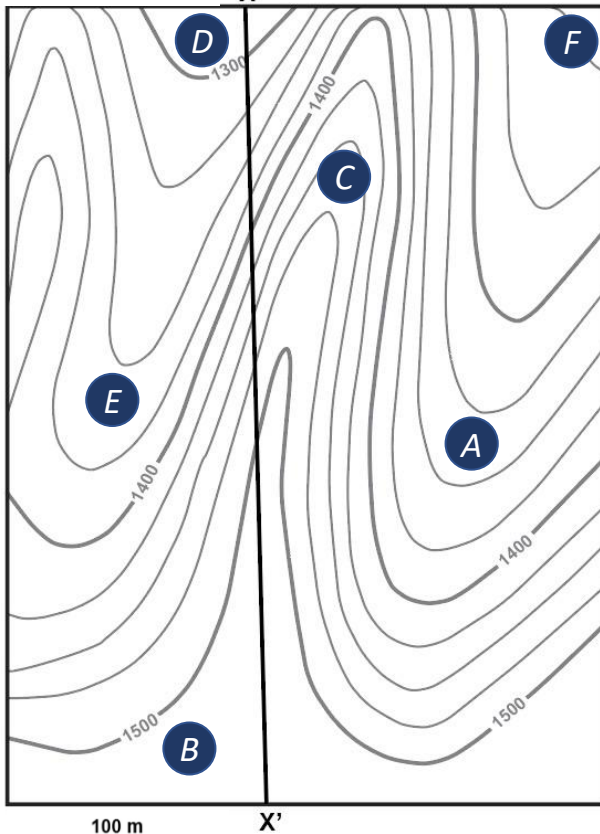
## EXERCISE ONE



**Learning Goal:** Identify contour lines and sketch topographic profiles from elevation data.

Examine the topographic map below and locate the contour interval. This number means that there are 25 meters of elevation between each contour line (in gray). First, align the wooden block with the surface shown below.

### Block 88D



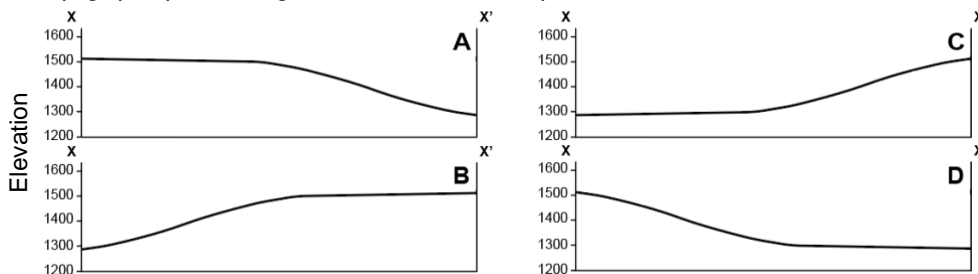
Contour Interval = 25 m

### Questions

- Locate where points A, B, C, D, E are on the topographic map and where they would approximately be on the wooden block model. Which point is located at the **highest** elevation? \_\_\_\_ Which point is located at the **lowest** elevation? \_\_\_\_ How do you know where the highest and lowest points are located?

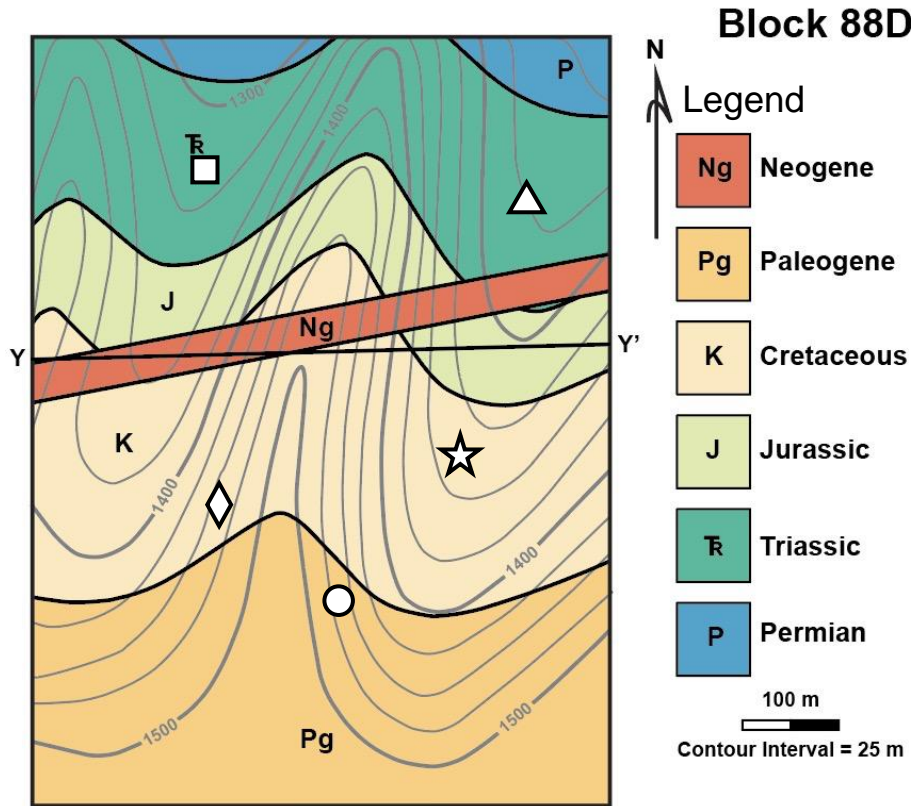
- Assuming there is no vegetation that could block your view and it is nighttime, if there was a campfire located at point E, at which of the other points would an observer be able to see the light from the fire? How do you know?

- Trace your finger along the path of X-X' on the map and the wooden block. Which one of the following is the correct topographic profile along the line from X at the top to X' at the bottom?



## EXERCISE ONE

The map below shows the same topography as the previous, but we have added a little more information. The legend now shows several rock units of various ages. In the legend, the units are ordered from youngest at the top to oldest at the bottom. Notice that there are shapes on this map to represent specific locations. Locate the square, triangle, star, diamond, and circle on the 2D map and its approximate location on the wooden block.



4. Assuming that there is no vegetation to block your view and it is nighttime, if there was a campfire located at the star, at which of the other positions (square, triangle, diamond, circle) would an observer be able to see the light from that fire? How do you know?

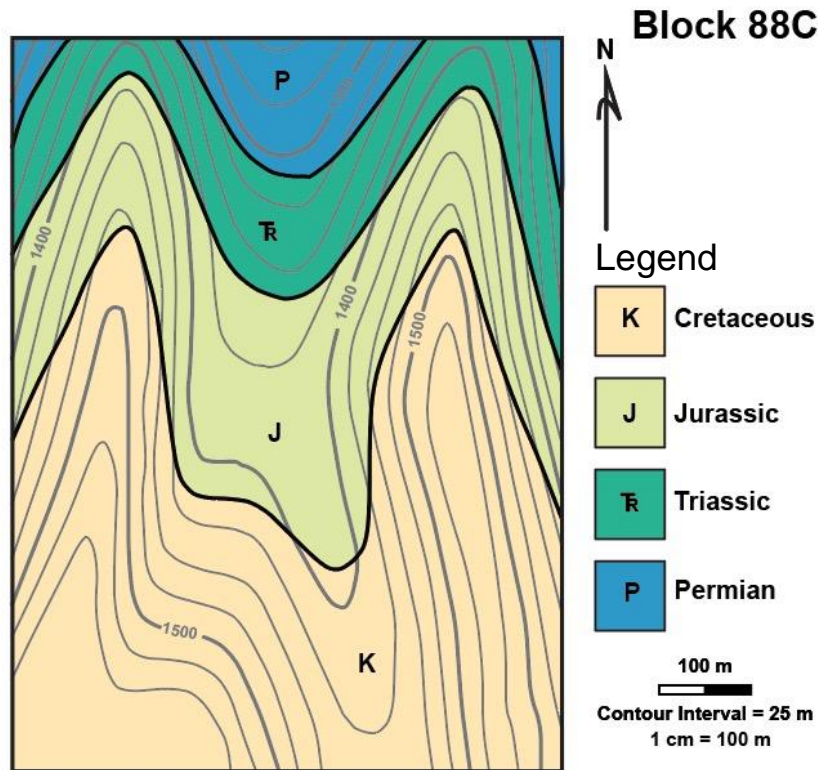


5. Trace your finger along the path of Y-Y' on the map and the wooden block. Sketch a topographic profile along the line from Y on the left to Y' on the right. Use the chart to the left to sketch the topographic profile.

## EXERCISE TWO

**Learning Goal:** Identify and differentiate between the contacts and faults on geologic maps.

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). Use your finger to trace the contacts on the wooden block.

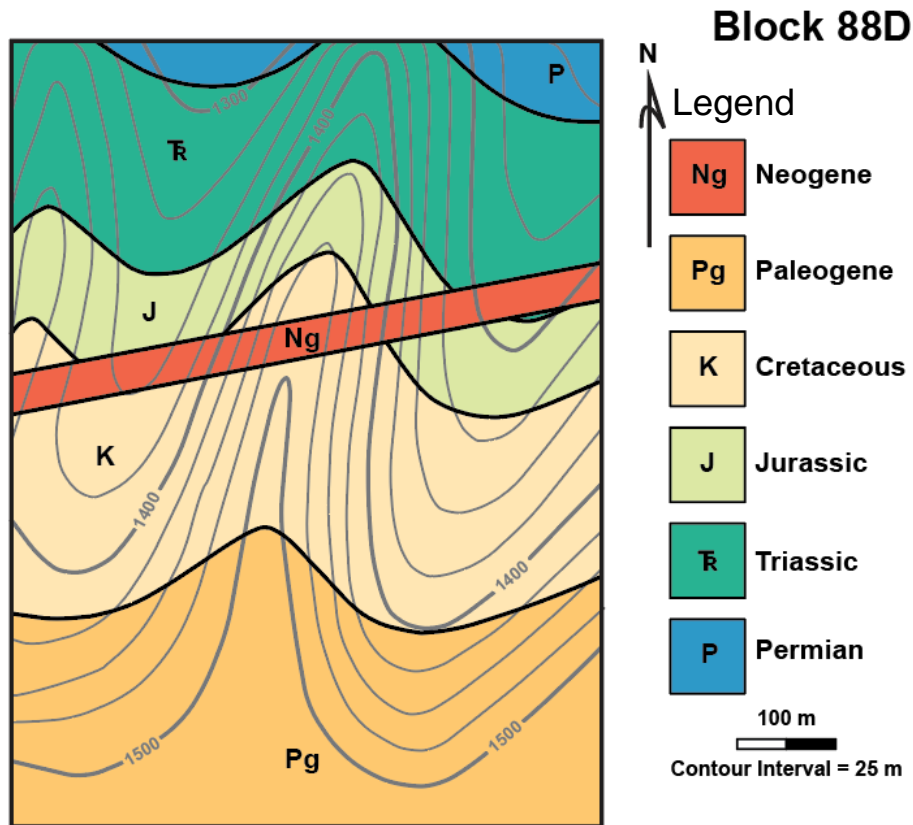


### Questions

1. Describe the appearances of the contour lines and the contacts. How would you describe their relationship?
2. By examining the contacts between the Jurassic and Triassic units, how steeply do these **units** dip? Shallow or steep? How do you know?
3. In what direction do these **units** dip? Use the North arrow provided with the map to help answer this question.

## EXERCISE TWO

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). Identify the feature of Neogene age that cuts across these beds. Trace this feature with your finger on the wooden block.



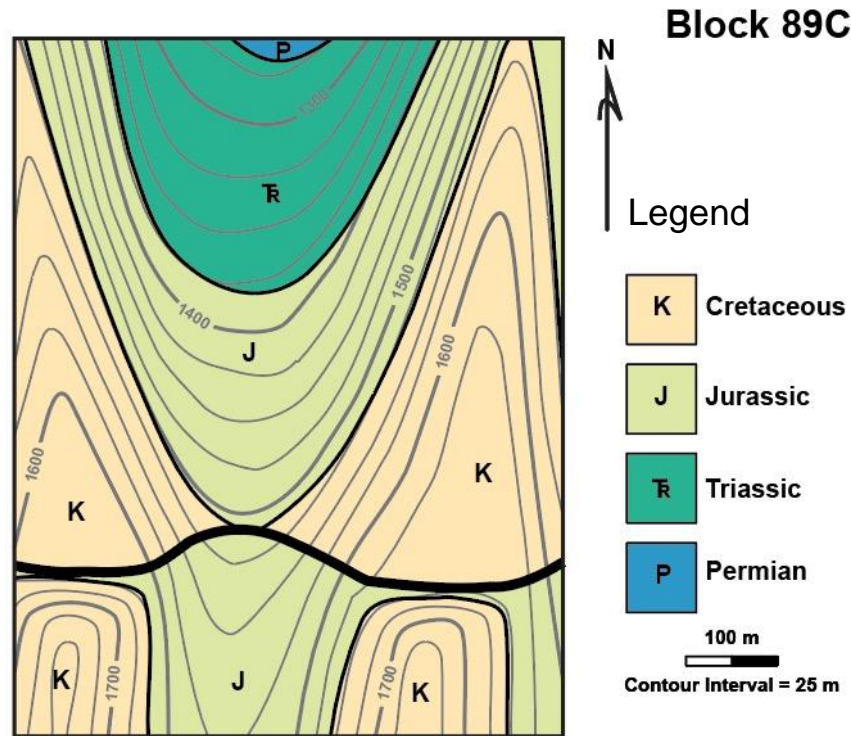
4. By examining the contacts between the Jurassic and Triassic units, how steeply do these **units** dip? Shallow or steep? How do you know?

5. There is a contact between the younger Neogene unit and the rest of the units on the map. Is this **contact** steeply or shallowly dipping?

### EXERCISE THREE

**Learning Goal:** Estimate the dip of the geologic units and structures and determine their direction of dip from geologic map data.

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). Now, we are adding a new thicker black line to indicate a fault. Label the fault on the map below. Use your finger to trace the **fault** on the wooden block. Identify the difference between the appearance of contacts and faults.

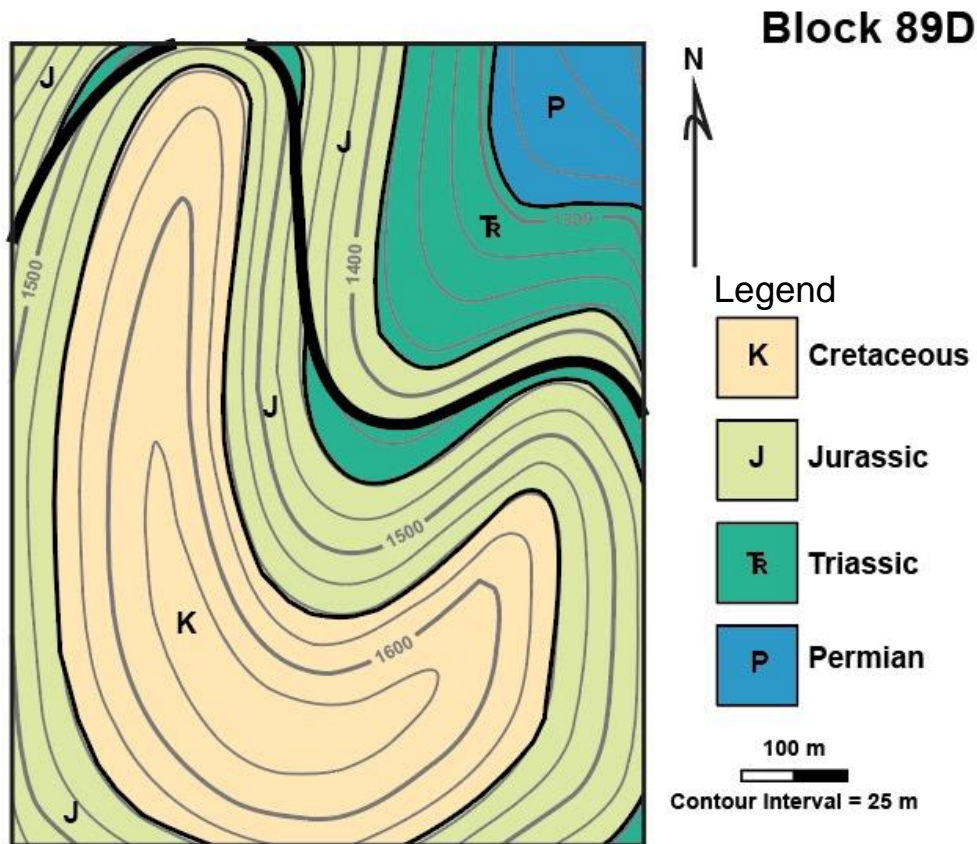


#### Questions

1. By examining the contacts between the Cretaceous and Jurassic units, how steeply do these **units** dip? Shallow or steep? How do you know?
2. Examine the fault and explain whether it is dipping shallowly or steeply.
3. What direction does the **fault** dip in? Use the North arrow on the map to help you answer.
4. What kind of fault is this? Normal, reverse/thrust, or strike-slip? How do you know?

### EXERCISE THREE

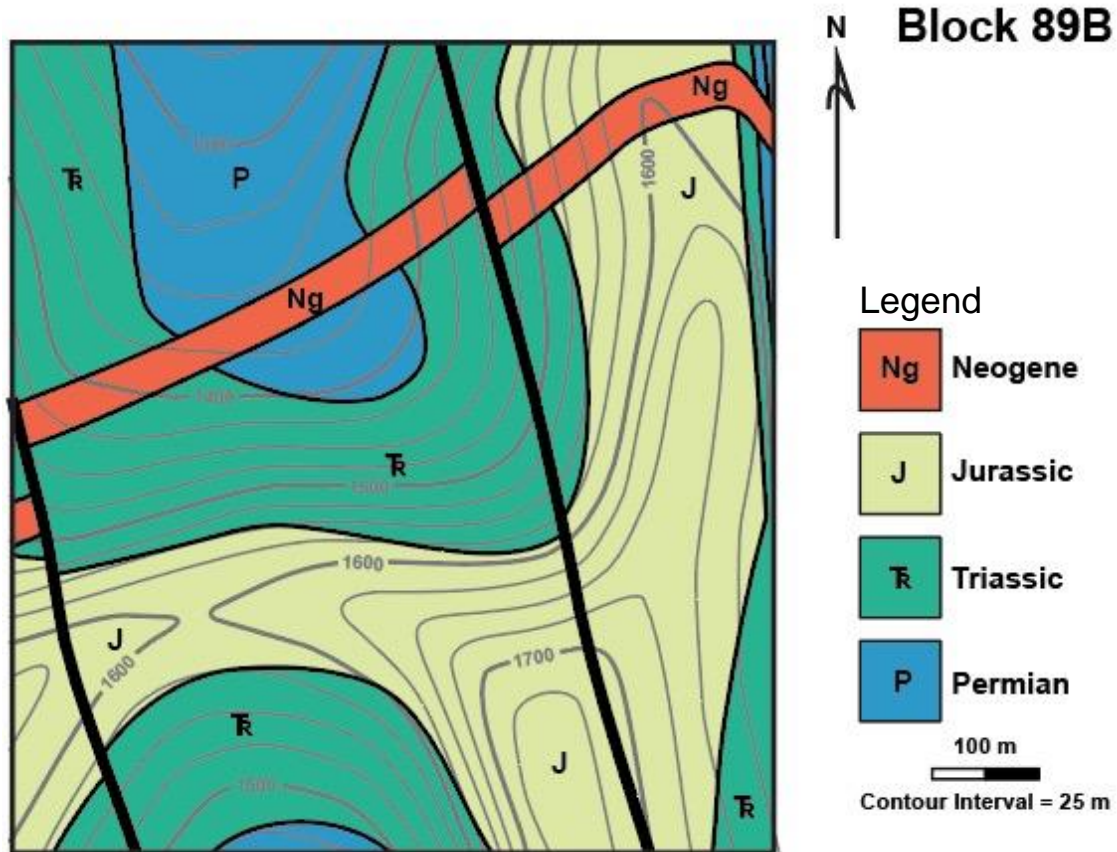
Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). The thicker black line indicates a fault. Label the fault on the map below.



5. By examining the contacts between the Cretaceous and Jurassic units, how steeply do these **units** dip? Shallow or steep? How do you know?
  
6. Examine the fault and explain whether it is dipping shallowly or steeply.
  
7. What direction does the **fault** dip in? Use the North arrow on the map to help you answer.
  
8. What kind of fault is this? Normal, reverse/thrust, or strike-slip? How do you know?

### EXERCISE THREE

Examine the map below and identify the contour lines (thinner, gray lines) and the contacts between geologic units (thicker, black lines). The thicker black line indicates a fault. Label the faults on the map below.



9. By examining the contacts between the Triassic and Jurassic units, how steeply do these **units** dip? Shallow or steep? How do you know?

10. Examine the fault and explain whether it is dipping shallowly or steeply.

11. What direction does the **fault** dip in? Use the North arrow on the map to help you answer.

12. What kind of fault is this? Normal, reverse/thrust, or strike-slip? How do you know?