Shaping Spatial Minds through Design Education

Linking Pupils' Spatial Thinking to Creative Ideation, Prototyping, and Making

Caiwei Zhu

Shaping Spatial Minds through Design Education

Linking Pupils' Spatial Thinking to Creative Ideation, Prototyping, and Making

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus,
Prof.dr.ir. H. Bijl,
chair of the Board for Doctorates
to be defended publicly on
Monday, 16 March 2026, 15:00

by

Caiwei ZHU

This dissertation has been approved by the (co)promotors.

Composition of the doctoral committee:

Rector Magnificus	chairperson
Prof.dr. M.J. de Vries,	Delft University of Technology, promotor
Dr.ir. R.M. Klapwijk,	Delft University of Technology, copromotor

Independent members:

Prof.dr.ir. C.R. Kleijn,	Delft University of Technology
Prof.dr. A.J. Cabo,	Delft University of Technology
Prof.dr. A.W. Lazonder,	Radboud University
Dr. M. Hickendorff,	Leiden University
Prof.dr. J.M.A.M. van Neerven,	Delft University of Technology, reserve member

This research was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Actions grant agreement No. 956124.

Printed by: Print Service Ede – The Netherlands
Cover by: Leonie Sonneveld
ISBN: 978-90-836732-3-3

An electronic version of this dissertation is available at: https://repository.tudelft.nl/

Shaping Spatial Minds through Design Education

Linking Pupils' Spatial Thinking to
Creative Ideation, Prototyping, and Making

Table of contents

Summary

Spatial thinking is a vital cognitive skill that spans and enriches learning across science, technology, engineering, and mathematics (STEM). Design, as an increasingly prevalent approach to STEM problem solving, provides rich opportunities to engage spatial thinking. However, limited research has examined how spatial thinking is enacted during open-ended design processes among primary and secondary school pupils, and findings regarding the relationship between spatial ability and design performance remain mixed. With the growing inclusion of Design and Technology (D&T) within K-12 curricula worldwide (with subject names that differ by country), the unique context of design education offers a valuable lens to deepen the field's emerging understanding of spatial thinking's role in design.

This dissertation investigates *what* forms of spatial thinking pupils engage in during design activities and *how* spatial ability relates to their creative design performance. Drawing on qualitative, mixed-methods, and quantitative approaches, this dissertation reports on four studies conducted in collaboration with international schools in the Netherlands, involving pupils aged 10 to 14.

Chapter 1 provides an overview of the importance of spatial thinking in STEM and reviews how educational practices in integrated learning contexts draw on and cultivate spatial skills. It highlights how design—both as a subject and as a practice that emphasizes solving real-world problems through an interdisciplinary approach—offers potential to support students' spatial developments.

Chapter 2 examines how pupils use spatial thinking in data physicalization activities that bridge STEM concepts and hands-on making. Visual-spatial representations, such as graphs, diagrams, and models, are crucial for effectively understanding and working with scientific, technological, engineering, and mathematical information. Data physicalization means transforming intangible data into tangible visual-spatial artifacts. While previous research has leveraged various forms of infographics to enrich learning and scaffold spatial skills development, little is known about how school-age pupils apply spatial thinking to design and create tangible data visualizations. This case study featured data physicalization activities conducted in two design classrooms at an international school in the Netherlands. By analyzing pupils' artifacts, classroom recordings, and semi-structured interviews, we identified seven themes related to how pupils constructed spatial representations of numerical information and developed both spatial and numerical understanding through this embodied data-physicalization making process.

Chapter 3 shifts the focus from broader STEM projects to a design project encompassing the design cycle—from exploring the design brief, developing and selecting design ideas, to prototyping with iterative refinement based on feedback and presenting the final solution. Although a growing body of research has examined spatial skills development in STEM learning and projects, many of these studies rely on closed-ended, analytical tasks. Consequently, there remains a limited understanding of the distinct spatial processes involved in open-ended design activities. Using a design-based research approach, we developed a nature-inspired, design-by-analogy project to investigate how young, novice designers engage in spatial thinking throughout the design process. Classroom observations and analyses of design artifacts collected at multiple stages revealed spatial thinking processes that appeared unique to open-ended, analogy-based

problem-solving processes that may not be captured by conventional spatial ability assessments. By illustrating the varied expressions of spatial thinking in creative contexts, this study offers practical insights for nurturing spatial and analogical reasoning in design education.

Drawing on observations from both case studies and emerging literature in the field, we observed that spatial thinking plays a meaningful role in STEM learning, particularly within design education. To extend these findings beyond the initial case studies, clarify the relationship between spatial ability and design performance, and address the limited and mixed evidence in existing research, we conducted a quantitative study examining the extent to which spatial ability contributes to pupils' creative design performance. This inquiry focused on creative design, with particular emphasis on the stages of design ideation and prototyping. These choices were guided by the recognition that creativity is an essential 21st-century competency and that design offers a rich context for exploring how creative ideas are generated and developed. While recent research has advanced understanding of spatial skills development through computer-aided or engineering design—often emphasizing later stages in design such as modeling and testing—less is known about the role of spatial thinking during earlier phases, which involves design ideation and initial prototyping. Accordingly, the third and fourth studies focused on the interplay between spatial thinking, creativity, and creative design performance during the early stages of the design process.

Chapter 4 explores the complexities of evaluating creativity in design ideas and prototypes. While D&T classrooms offer ample opportunities for fostering creativity, the open-ended and context-dependent nature of design problems often complicates the evaluation of pupils' creative work. Comparative judgement, an assessment method that relies on holistic evaluations by expert judges rather than predefined criteria, provides a promising alternative by capturing qualities that traditional rubric-based assessments might overlook. In this study, 20 industrial design master's students or recent graduates evaluated the creativity of design ideas and prototypes produced by 201 pupils. By combining qualitative and quantitative analyses, the findings contribute to a more nuanced understanding of creativity assessment in design education contexts.

Building on the holistic evaluation, Chapter 5 examines whether and how pupils' spatial ability is related to their creative design performance. Prior research has reported mixed findings, with some studies indicating positive associations between spatial ability and creative design performance among university students; however, its predictive role among younger learners and its relative importance compared to other cognitive skills remain unclear. The final study investigated how spatial ability, divergent thinking, and creative imagery predict design performance across two stages—ideation and prototyping—in 201 pupils aged 10 to 14. While spatial ability correlated with both ideation and prototype quality, hierarchical regression analyses revealed a nuanced pattern: after controlling for age and verbal ability, creative imagery predicted ideation quality, whereas spatial ability did not. Conversely, spatial ability emerged as a significant predictor of prototype quality, representing a unique contribution independent of divergent thinking and creative imagery. These results suggest that spatial ability plays a distinct role when pupils develop their initial ideas into novel and functionally viable design prototypes.

Taken together, this dissertation characterizes the diverse ways in which pupils engage in spatial thinking during design and examines the extent to which their spatial ability contributes to creative design performance. The first two studies provide detailed insights into how pupils apply spatial thinking during design and maker practices. When making data physicalization, pupils leveraged the visual and spatial properties of self-selected materials to encode numerical data, including mapping quantities to sizes, unitizing materials proportionally, and representing ordinal relationships spatially. This hands-on making process prompted pupils to experiment with ways to give visual-spatial forms to data, reflect on their ability to craft 3D artifacts, and consider how their artifacts would be perceived from different angles. When developing nature-inspired designs, pupils abstracted visual-spatial features from source examples to infer form–function relationships, and subsequently generated diverse analogical matches. Across both contexts, pupils employed spatial thinking at various stages throughout the design process, from generating and elaborating ideas to visualizing how their design prototypes would function in real-world scenarios.

Observations from the design classrooms revealed notable distinctions between spatial thinking in design practice and the ways spatial ability is typically addressed in standardized assessments. In design contexts, pupils reconfigure and transform known information into artefacts that are novel and functional. This process involves mentally exploring and manipulating visual-spatial information, as well as visually or physically representing ideas through prototyping in order to identify potential improvements in visual-spatial or functional features. In contrast, spatial tests primarily tap into the ability to form, retain, and manipulate mental representations in order to identify pre-determined answers, with limited emphasis on generating original solutions or creative refinement. Thus, the spatial thinking needed for solving open-ended design problems may differ from that involved in solving closed-ended spatial tests in terms of their goals and processes.

The quantitative analyses clarified how psychometrically assessed spatial ability contributes to pupils' creative design performance. Notably, the finding that spatial ability significantly predicted design prototype quality—but not ideation—after controlling for pupils' age, verbal ability, and creativity underscores its unique role in transforming "out-of-the-box" ideas into novel and viable prototypes. Although the examined variables only explained a small portion of the variance in pupils' creative design performance, these results provide some of the first empirical evidence of the relationship between spatial ability and creative design performance among K-12 pupils. These findings point to the need for future studies to examine additional factors influencing design performance and to develop design-specific pedagogies that effectively scaffold pupils' spatial skills development.

This dissertation offers several recommendations for future research and for the purposeful embedding of spatial practices in classrooms. Prior qualitative evidence from both design studies and design education research consistently highlights the importance of visual-spatial skills in design practice; however, quantitative investigations have yielded mixed findings regarding the relationship between spatial ability and design performance. This discrepancy suggests potential limitations in the formulation and applicability of current psychometric measures of spatial ability, as well as complexities in evaluating creative design performance. These observations, together with findings of this dissertation, highlight the need for further research to examine how the choice of spatial instruments (domain-general versus domain-specific) and the nature of learning tasks (open-ended versus closed-ended) influence the relationship between spatial ability and subject-specific performance.

In alignment with current spatial ability research and the findings uncovered in this dissertation, it is recommended that pupils' spatial skills development be integrated with the practical goal of enabling them to effectively solve real-world STEM problems. With a growing body of research pointing toward the integration of spatial training within curricula, this dissertation recommends fostering spatial skills specifically in interdisciplinary and transdisciplinary STEM contexts to ensure that training reflects the scenarios in which these skills are applied. Given that design thinking and practices are key tools across STEM disciplines, and that both the qualitative and quantitative findings of this dissertation highlight spatial ability's meaningful role in design, educators are encouraged to leverage design-based STEM challenges. Such challenges have the potential to cultivate spatial skills while strengthening pupils' capacity to tackle authentic and complex problems, thereby creating a mutually reinforcing cycle of cognitive development and practical application.

Samenvatting

Ruimtelijk inzicht is een essentiële mentale vaardigheid die het leerproces binnen de bèta wetenschap (in het Engels: ‘STEM’) omvat en verrijkt. Ontwerpen, als een steeds vaker voorkomende aanpak van probleemoplossing binnen STEM, biedt rijke mogelijkheden om ruimtelijk inzicht te stimuleren. Er is echter beperkt onderzoek gedaan naar hoe ruimtelijk inzicht wordt toegepast tijdens vrijblijvende ontwerpprocessen bij leerlingen in het primair en voortgezet onderwijs, en de bevindingen over het verband tussen ruimtelijk inzicht en ontwerpprestaties zijn nog steeds tegenstrijdig. Met de toenemende integratie van Design and Technology (D&T) in het curriculum van het primair en voortgezet onderwijs over de hele wereld (met vaknamen die vaak per land verschillen), biedt de unieke context van ontwerponderwijs een waardevol perspectief om het groeiende begrip van de rol van ruimtelijk inzicht in het ontwerpproces te verdiepen.

Dit proefschrift onderzoekt welke vormen van ruimtelijk inzicht leerlingen toepassen tijdens ontwerpactiviteiten en hoe ruimtelijk inzicht samenhangt met hun creatieve ontwerpprestaties. Dit proefschrift maakt gebruik van een kwalitatieve, gemengde en kwantitatieve aanpak en beschrijft vier onderzoeken die in samenwerking met internationale scholen in Nederland zijn uitgevoerd met leerlingen van 10 tot 14 jaar.

Hoofdstuk 1 geeft een overzicht van het belang van ruimtelijk inzicht in de bètavakken (Eng.: STEM) en beschrijft hoe onderwijspraktijken in een geïntegreerde leeromgeving ruimtelijke vaardigheden benutten en ontwikkelen. Het benadrukt hoe ontwerp – zowel als vak als in de praktijk waar de nadruk ligt op het oplossen van alledaagse problemen doormiddel van een interdisciplinaire aanpak – de potentie heeft om de ruimtelijke ontwikkeling van leerlingen te ondersteunen.

Hoofdstuk 2 onderzoekt hoe leerlingen ruimtelijk inzicht gebruiken in activiteiten gericht op data-fysicalisatie, die een brug slaan tussen technische concepten en praktische toepassingen. Visueel-ruimtelijke figuren, zoals grafieken, diagrammen en prototypes, zijn cruciaal voor het effectief begrijpen en verwerken van wetenschappelijke, technische en wiskundige informatie. Data-fysicalisatie betekent het transformeren van abstracte data in tastbare visueel-ruimtelijke artefacten. Hoewel eerder onderzoek verschillende vormen van infographics heeft ingezet om het leerproces te verbeteren en de ontwikkeling van ruimtelijke vaardigheden te ondersteunen, is er weinig bekend over hoe scholieren ruimtelijk inzicht toepassen in het ontwerpproces en bij het creëren van fysieke datavisualisaties. Deze casestudy beschreef activiteiten voor het fysiek visualiseren van data in twee ontwerpklassen op een internationale school in Nederland. Door de prototypes van de leerlingen, opnames van de lessen en semi-gestructureerde interviews te analyseren, identificeerden we zeven thema's die verband hielden met de manier waarop leerlingen ruimtelijke representaties van numerieke informatie construeerden en zowel ruimtelijk als numeriek begrip ontwikkelden doormiddel van dit concrete proces van datavisualisatie.

Hoofdstuk 3 verschuift de focus van bredere techniekprojecten naar een ontwerpproject dat de ontwerpcyclus omvat – van het verkennen van de ontwerpopdracht, het ontwikkelen en selecteren van ontwerpideeën, tot het prototypen met iteratieve verfijning op basis van feedback en het presenteren van de uiteindelijke oplossing. Hoewel er veel onderzoek is gedaan naar de ontwikkeling van ruimtelijke vaardigheden in STEM-onderwijs en -projecten, maken veel van deze studies gebruik van vaste, analytische taken.

Dit heeft als een gevolg dat er een beperkt begrip is van de specifieke ruimtelijke processen die betrokken zijn bij open ontwerpactiviteiten. Met behulp van een ontwerpgerichte onderzoeksaanpak ontwikkelden we een op de natuur geïnspireerd ontwerp-door-analogieproject om te onderzoeken hoe jonge, beginnende ontwerpers ruimtelijk inzicht toepassen gedurende het ontwerpproces. Observaties in de klas en analyses van ontwerpobjecten die in verschillende fasen werden verzameld, onthulden ruimtelijke denkprocessen die uniek leken voor open, op analogie gebaseerde probleemoplossingsprocessen en die mogelijk niet worden vastgelegd door gebruikelijke beoordelingen van ruimtelijk inzicht. Door de diverse uitingen van ruimtelijk inzicht in de creatieve context te illustreren, biedt deze studie praktische inzichten voor het stimuleren van ruimtelijk en analogisch redeneren in het ontwerponderwijs.

Op basis van observaties uit zowel casestudies als recente literatuur in het vakgebied, hebben we vastgesteld dat ruimtelijk inzicht een belangrijke rol speelt in het bèta-onderwijs, met name binnen het ontwerponderwijs. Om deze bevindingen verder uit te breiden dan de eerste aanvankelijke casestudies, de relatie tussen ruimtelijk inzicht en ontwerpprestaties te verduidelijken en de beperkte en tegenstrijdige resultaten in bestaand onderzoek aan te pakken, hebben we een kwantitatieve studie uitgevoerd naar de mate waarin ruimtelijk inzicht bijdraagt aan de creatieve ontwerpprestaties van leerlingen. Dit onderzoek richtte zich op creatief ontwerp, met name op de fasen van ideeënvorming en prototyping. Deze keuzes werden ingegeven door de erkenning dat creativiteit een essentiële vaardigheid is voor de moderne tijd en dat ontwerp een rijke context biedt voor het onderzoeken van hoe creatieve ideeën worden gegenereerd en ontwikkeld. Hoewel recent onderzoek het begrip van de ontwikkeling van ruimtelijke vaardigheden doormiddel van computerondersteund of technisch ontwerp heeft vergroot – vaak met de nadruk op latere fasen in het ontwerpproces zoals modelleren en testen – is er minder bekend over de rol van ruimtelijk inzicht tijdens eerdere fasen, zoals ideeënvorming en de aanvankelijke prototyping. Daarom richtten het derde en vierde onderzoek zich op de wisselwerking tussen ruimtelijk inzicht, creativiteit en creatieve ontwerpprestaties tijdens de vroege stadia van het ontwerpproces.

Hoofdstuk 4 onderzoekt de complexiteit van het beoordelen van creativiteit in ontwerpideeën en prototypes. Hoewel ontwerp- en technieklessen volop mogelijkheden bieden om creativiteit te stimuleren, bemoeilijkt het open karakter en de contextafhankelijkheid van ontwerpproblemen vaak de beoordeling van het creatieve werk van leerlingen. Vergelijkende beoordeling, een beoordelingsmethode die gebaseerd is op holistische evaluaties door deskundige juryleden in plaats van vooraf vastgestelde criteria, biedt een veelbelovend alternatief door kwaliteiten vast te leggen die traditionele, op rubrieken gebaseerde beoordelingen mogelijk over het hoofd zien. In deze studie beoordeelden 20 masterstudenten industrieel ontwerp of recent afgestudeerden de creativiteit van ontwerpideeën en prototypes die door 201 leerlingen waren geproduceerd. Door kwalitatieve en kwantitatieve analyses te combineren, dragen de bevindingen bij aan een genuanceerder begrip van creativiteitsbeoordeling in de context van ontwerponderwijs.

Voortbouwend op de holistische evaluatie onderzoekt hoofdstuk 5 of, en hoe, het ruimtelijk inzicht van leerlingen samenhangt met hun prestaties op het gebied van creatief ontwerpen. Eerder onderzoek heeft wisselende resultaten laten zien, waarbij sommige studies een positieve associatie aantonen tussen ruimtelijk inzicht en creatieve ontwerpprestaties bij universiteitsstudenten. De voorspellende rol ervan bij jongere leerlingen en het relatieve belang ervan ten opzichte van andere cognitieve vaardigheden blijven echter onduidelijk. De laatste studie onderzocht hoe ruimtelijk inzicht, divergent denken en creatieve beeldvorming de ontwerpprestaties voorspellen in twee fasen – ideeënvorming en prototyping – bij 201 leerlingen van 10 tot 14 jaar. Hoewel ruimtelijk inzicht correleerde met zowel de kwaliteit van de ideeënvorming als de prototypekwaliteit, onthulden hiërarchische regressieanalyses een genuanceerd patroon: na correctie voor leeftijd en verbale vaardigheden voorspelde creatieve beeldvorming de kwaliteit van de ideeënvorming, terwijl ruimtelijk inzicht dat niet deed. Omgekeerd bleek ruimtelijk inzicht een significante voorspeller te zijn van de prototypekwaliteit, wat een unieke bijdrage vertegenwoordigt, onafhankelijk van divergent denken en creatieve beeldvorming. Deze resultaten suggereren dat ruimtelijk inzicht een aparte rol speelt wanneer leerlingen hun eerste ideeën ontwikkelen tot nieuwe en functioneel bruikbare ontwerpprototypes.

Samengevat beschrijft dit proefschrift de diverse manieren waarop leerlingen ruimtelijk inzicht toepassen tijdens het ontwerpen en onderzoekt het in hoeverre hun ruimtelijk inzicht bijdraagt aan creatieve ontwerpprestaties. De eerste twee studies bieden gedetailleerde inzichten in hoe leerlingen ruimtelijk inzicht toepassen tijdens ontwerp- en maakprocessen. Bij het fysiek visualiseren van data maakten leerlingen gebruik van de visuele en ruimtelijke eigenschappen van zelfgekozen materialen om numerieke data te coderen, waaronder het koppelen van hoeveelheden aan groottes, het proportioneel verdelen van materialen en het ruimtelijk weergeven van ordinale relaties. Dit praktische maakproces zette leerlingen ertoe aan te experimenteren met manieren om data visueel-ruimtelijke vormen te geven, te reflecteren op hun vermogen om 3D-objecten te maken en na te denken over hoe hun objecten vanuit verschillende hoeken zouden worden waargenomen. Bij het ontwikkelen van op de natuur geïnspireerde ontwerpen abstraheerden leerlingen visueel-ruimtelijke kenmerken uit bronvoorbeelden om vorm-functie-relaties af te leiden en genereerden vervolgens diverse analoge overeenkomsten. In beide contexten pasten leerlingen ruimtelijk inzicht toe in verschillende fasen van het ontwerpproces, van het genereren en uitwerken van ideeën tot het visualiseren van hoe hun ontwerpprototypes zouden functioneren in realistische scenario's.

Observaties in de ontwerplessen brachten opmerkelijke verschillen aan het licht tussen ruimtelijk inzicht in de ontwerppraktijk en de manier waarop ruimtelijk inzicht doorgaans wordt getoetst in gestandaardiseerde beoordelingen. In ontwerpcontexten herconfigureren en transformeren leerlingen bekende informatie tot nieuwe en functionele objecten. Dit proces omvat het mentaal verkennen en manipuleren van visueel-ruimtelijke informatie, evenals het visueel of fysiek weergeven van ideeën doormiddel van het maken van prototypes om potentiële verbeteringen in visueel-ruimtelijke of functionele kenmerken te identificeren. Ruimtelijke toetsen daarentegen controleren voornamelijk het vermogen om mentale representaties te vormen, te onthouden en te manipuleren om vooraf bepaalde antwoorden te vinden, met beperkte nadruk op het genereren van originele oplossingen of creatieve verfijning. Het ruimtelijk inzicht dat nodig is voor het oplossen van vrijblijvende ontwerpproblemen kan dus verschillen van het denken dat nodig is voor het oplossen van gesloten ruimtelijke toetsen, zowel wat betreft de doelen als de processen.

De kwantitatieve analyses hebben duidelijk gemaakt hoe psychometrisch vastgesteld ruimtelijk inzicht bijdraagt aan de creatieve ontwerpprestaties van leerlingen. Met name de bevinding dat ruimtelijk inzicht de kwaliteit van ontwerpprototypes – maar niet de ideeëngeneratie – significant voorspelde na correctie voor de leeftijd, verbale vaardigheden en creativiteit van de leerlingen. Dit onderstreept de unieke rol van ruimtelijk inzicht bij het omzetten van originele ideeën in nieuwe en haalbare prototypes.

Hoewel de onderzochte variabelen slechts een klein deel van de variantie in de creatieve ontwerpprestaties van leerlingen verklaarden, bieden deze resultaten een van de eerste empirische bewijzen voor de relatie tussen ruimtelijk inzicht en creatieve ontwerpprestaties bij leerlingen in het primair en voortgezet onderwijs. Deze bevindingen wijzen op de noodzaak van toekomstig onderzoek naar aanvullende factoren die de ontwerpprestaties beïnvloeden en naar de ontwikkeling van ontwerpspecifieke pedagogische methoden die de ontwikkeling van ruimtelijke vaardigheden bij leerlingen effectief ondersteunen.

Dit proefschrift biedt diverse aanbevelingen voor toekomstig onderzoek en voor de doelgerichte integratie van ruimtelijke vaardigheden in het onderwijs. Voorgaand kwalitatief onderzoek, zowel in ontwerpstudies als in ontwerponderwijs, benadrukt consistent het belang van visueel-ruimtelijke vaardigheden in de ontwerppraktijk. Kwantitatief onderzoek heeft echter wisselende resultaten opgeleverd met betrekking tot de relatie tussen ruimtelijk inzicht en ontwerpprestaties. Deze discrepantie wijst op mogelijke beperkingen in de formulering en toepasbaarheid van de huidige psychometrische meetinstrumenten voor ruimtelijk inzicht, evenals op de complexiteit van het evalueren van creatieve ontwerpprestaties. Deze observaties, samen met de bevindingen van dit proefschrift, benadrukken de noodzaak van verder onderzoek naar de invloed van de keuze van ruimtelijke instrumenten (domein-algemeen versus domein-specifiek) en de aard van de leertaken (open versus gesloten vragen) op de relatie tussen ruimtelijk inzicht en vakspecifieke prestaties.

In lijn met huidig onderzoek naar ruimtelijk inzicht en de bevindingen van dit proefschrift, wordt aanbevolen de ontwikkeling van ruimtelijke vaardigheden bij leerlingen te integreren met het praktische doel hen in staat te stellen effectief realistische technische problemen (Eng.: STEM) op te lossen. Gezien het groeiende aantal onderzoeken dat wijst op de integratie van ruimtelijke training in curricula, beveelt dit proefschrift aan om ruimtelijke vaardigheden specifiek te bevorderen in interdisciplinaire en transdisciplinaire technische contexten, zodat de training aansluit op de scenario's waarin deze vaardigheden worden toegepast. Aangezien ontwerppraktijken en gedachtes belangrijke instrumenten zijn binnen alle bèta-disciplines, en zowel de kwalitatieve als kwantitatieve bevindingen van dit proefschrift de betekenisvolle rol van ruimtelijk inzicht in ontwerp benadrukken, worden docenten aangemoedigd om ontwerpgerichte techniekuitdagingen te benutten. Dergelijke uitdagingen kunnen ruimtelijke vaardigheden ontwikkelen en tegelijkertijd het vermogen van leerlingen versterken om authentieke en complexe problemen aan te pakken, waardoor een wederzijds versterkende cyclus van cognitieve ontwikkeling en praktische toepassing ontstaat.

CHAPTER 1
Spatial thinking in STEM and design education

1.1 Spatial thinking in STEM education[1]

Much of contemporary society values scientific knowledge and technological development as key contributors to economic growth, innovation, and welfare (Davies & Horst, 2016). Attracting more students to Science, Technology, Engineering, and Mathematics (STEM) at the primary, secondary, and tertiary levels of education has been on national agendas across the world (Gough, 2015; National Research Council, 2011). Variables such as interest in STEM (Caprile et al., 2015), sense of belonging in STEM fields (Dortch & Patel, 2017), and self-efficacy in learning STEM (Tracey, 2010) influence students' decisions to pursue STEM studies and careers, and have been the focus of initiatives aimed at increasing STEM participation (Caprile et al., 2015). Comparatively less attention has been given to cognitive factors, such as students' spatial ability, despite evidence that spatial ability is a key predictor of academic and professional involvement in STEM fields (Kell et al., 2013; Shea et al., 2001; Wai et al., 2009).

Spatial thinking is ubiquitous in everyday life. We use spatial thinking when estimating how much paper is needed to wrap a present, following instructions to assemble flat-pack furniture, or determining whether we are standing on the platform for the bus traveling in the direction we need to go. Essentially, spatial thinking refers to the processes of understanding spatial concepts, using mental and external representations, and reasoning about relationships between objects within a given space and relative to that space (National Research Council, 2006). Spatial ability, with a tradition rooted in educational and psychological testing, can be broadly understood as "the ability to generate, retain, retrieve, and transform well-structured visual images" (Lohman, 1996, p. 112). Such ability is not confined to two-dimensional representations but also involves making sense and use of spatial information embedded in three-dimensional objects and environments (Carroll, 1993; Ramey & Uttal, 2017). It encompasses various subskills, such as mental rotation, spatial visualization, and spatial orientation, each of which contributes to solving problems that involve spatial relationships.

Over the past two decades, a growing body of research has been devoted to investigating the relationship between students' spatial ability and academic performance, as well as training approaches aimed at improving spatial ability. Spatial ability consistently and meaningfully predicts performances in sciences (Harle & Towns, 2011; Hodgkiss et al., 2018), technology (Bockmon et al., 2020; Chang, 2014), engineering (Sorby, 2009; Duffy et al., 2020), and mathematics (Atit et al., 2022; Mix et al., 2016). Although much research on spatial ability and STEM performance has focused on higher education (e.g., Sorby, 2009; Hegarty, 2014), it is critical to make use of the formal and informal educational resources from early childhood on (Newcombe, 2010; Newcombe & Frick, 2010). Research on mathematics learning, for instance, shows that spatial ability is positively correlated with math performance among preschoolers (Rasmussen & Bisanz, 2005), primary school pupils (Lowrie et al., 2016), as well as secondary school students and beyond (Webb et al., 2007).

In addition to its importance for STEM outcomes, spatial ability can be developed

[1] *This chapter includes text that has previously appeared, with modifications, in the following source:*

Zhu, C., Leung, C. O-Y., Lagoudaki, E., Velho, M., Segura-Caballero, N., Jolles, D., Duffy, G., Maresch, G., Pagkratidou, M., & Klapwijk, R. (2023). Fostering spatial ability development in and for authentic STEM learning. *Frontiers in Education, 8,* 1138607.

through educational training. A meta-analysis of over 200 studies on various forms of spatial training—including direct spatial training, informal training such as learning through games and play, and spatial training embedded in curricula—found that spatial training consistently yielded significant gains in spatial task performance for both males and females (Uttal, Meadow, et al., 2013). These gains are practically valuable when they result in enhanced STEM performance (Sorby et al., 2018; Stieff & Uttal, 2015). To date, research from multiple disciplines has been conducted to foster spatial ability development in K-12 educational contexts, including cognitive psychology (e.g., Uttal, Miller & Newcombe, 2013; Hawes et al., 2017; Mix, 2019; Gilligan et al., 2020), educational science (e.g., Sorby, 2009; Newcombe & Frick, 2010; Casey et al., 2011; Lowrie et al., 2019), and STEM education (e.g., Burte et al., 2017; Ramey & Uttal, 2017; Julià & Antolì, 2018; Atit et al., 2020).

Among spatial interventions delivered in physical, digital, or hybrid formats, as well as efforts to spatialize STEM curricula, a growing trend is to align training materials and modes more closely with students' everyday learning experiences. Researchers have underscored the need for spatial training that authentically addresses the spatial skills students use to solve real-world STEM problems (e.g., Ormand et al., 2014; Atit et al., 2020; Ramey et al., 2020). Whether through the use of hands-on manipulatives, digital feedback on spatial tasks, or the adaptation of student-centered pedagogies to support spatial development, such approaches reflect the intentional recruitment of authentic, contextually relevant teaching and learning methods. Authentic learning means "situating learning tasks in the context of future use" and in "real-world situations" (Herrington et al., 2014, p. 401). To enhance students' spatial ability in authentic learning contexts, we need to prepare students with spatial knowledge and spatial thinking skills that address discipline-specific challenges (Hinze et al., 2014; Stieff et al., 2014; Atit et al., 2020). In this regard, integrated STEM education—which encompasses the interdisciplinary essence of STEM and the interplay among scientific inquiry, technological literacy, engineering design, and mathematical thinking (Kelley & Knowles, 2016)—offers promising opportunities to support K-12 students' spatial development through authentic learning experiences.

1.2 Fostering spatial skills development through integrated STEM with design as a bridge

STEM education is, at its core, an intersection of different disciplines. Instead of describing STEM as merely a cluster of science, technology, engineering, and mathematics subjects, the concept of integrated STEM education has gained increasing recognition from educational researchers in North America, Europe, and Asia (e.g., Bryan et al., 2015; Mustafa et al., 2016; Stohlmann et al., 2012). Kelley and Knowles (2016) defined integrated STEM education as "the approach to teaching the STEM content of two or more STEM domains, bound by STEM practices within an authentic context for the purpose of connecting these subjects to enhance student learning" (p. 3). Carefully designed integrated STEM education can lead to a range of desirable learning outcomes, such as increased knowledge in STEM (Kelley et al., 2023), improved test performance in subjects such as math and science (Tillman et al., 2014), enhanced problem-solving skills (Netwong, 2018), a more positive attitude toward STEM (Sisman et al., 2021), and higher levels of engagement with STEM courses (Taylor & Hutton, 2013; Peng & Sollervall, 2014).

Because integrated STEM education situates disciplinary content within interdisciplinary problem-solving contexts, it also provides a promising environment for supporting students' spatial development. Instead of treating spatial skills as a separate domain, the National Research Council in the U.S. recommended viewing it as a "missing link" across different subject knowledge that "permeates" different disciplines (2006, p. 7). Thus, understanding how to develop students' spatial thinking skills in integrated STEM education not only aligns with the goal of developing spatial skills alongside content knowledge but also renders a possible answer to how spatial thinking may connect learning across different STEM disciplines.

Within their conceptual framework for integrated STEM education, Kelley and Knowles (2016) highlighted four learning approaches that well accommodate integrated STEM learning, including scientific inquiry, technological literacy, engineering design, and mathematical thinking. These approaches are grounded in shared practices across disciplines and the situated context for learning, so that "learning is authentic and relevant, therefore representative of an experience found in actual STEM practice" (p. 4). In the following subsections, we present an overview of how existing spatial interventions or spatialized educational practices map onto each of the four integrated STEM approaches, and how integrated STEM problems that originate from real-world problems can be used to develop students' spatial skills. We further examine how design, acting as a bridge that weaves knowledge and skills practiced in integrated STEM projects, offers still-underexplored opportunities to deepen the understanding of, and further develop, spatial skills.

1.2.1 Engaging students in spatial thinking through scientific inquiry

Learning science through inquiry is essential to prepare students for scientific investigations. Existing inquiry-based educational practices have shown an effort to address students' spatial ability development. Interpreting and creating scientific diagrams, for example, demand spatial thinking (Newcombe, 2010). In secondary chemistry classrooms, Stieff (2011) investigated the use of computer-based visualization in a guided-inquiry curriculum to develop students' competence in working with scientific representations. By exploring and visualizing the properties of different chemical substances and the dynamic chemical processes, students in the visualization-focused curriculum developed a better understanding of the content knowledge than those who received traditional lecture instruction. They also demonstrated more competency in creating scientific representations like chemistry professionals do.

For another example, Oberle (2020) recorded sixth, seventh, and eighth-grade students' learning through the National Geographic Geo-Inquiry Process, a curriculum that presents rich opportunities for spatial thinking, including comprehending geographic representations and creating geographic representations using data or maps. While this study did not explicitly examine gains in spatial skills among the students, the learning of geographic skills inherently demands multiple types of spatial thinking, such as spatial orientation, spatial visualization, and complex spatial reasoning (National Research Council, 2006). Compared to the control group who received a traditional, non-inquiry-focused curriculum, those who participated in the Geo-Inquiry class showed modest improvement in tasks such as discussing spatial patterns at different scales and elaborating on spatial patterns using maps.

1.2.2 Engaging students in spatial thinking through mathematical thinking

Cohrssen et al. (2017) and Cohrssen and Pearn (2021) incorporated project-based learning, which is another important way to practice integrated STEM (Ritz & Fan, 2015), to teach soon-to-be elementary school children about spatial thinking and consequently support their mathematical thinking. Children who were facing a transition from kindergarten to elementary school were challenged to learn the route to their new school and visually represent their route through map-making. The core and complementary activities provided ample opportunities for children to actively reason about spatial orientation and spatial visualization. Moreover, children were encouraged to use directional and locational language and other symbolic representations, such as gesturing and sketching, to represent their thinking. Taking a qualitative lens, these studies allowed researchers to dive deep into how children's spatial thinking has been developed during these activities. For example, providing children with the vocabulary to describe positions, directions, and 2D and 3D shapes, together with prompting children's use of such words through questions, led to fruitful spatial and mathematical conversations between children and teachers. In addition to exercising navigation skills, using spatial memory, and making spatial representations, these children were also actively using their mathematics knowledge during the project-based learning experience.

1.2.3 Engaging students in spatial thinking through technological literacy

Technology, as Kelley and Knowles phrased, is not just an important tool or vehicle through which students learn about science, engineering, and mathematics, but is itself a discipline consisting of knowledge and practices involved in "designing, making, and using" of technology (p. 6). As the use of technology is ubiquitous across various disciplines, opportunities exist to spatialize technology-focused or technology-enhanced learning experiences. For example, Bhaduri et al. (2021) conducted a pilot study to teach seventh and eighth-grade students 3D modeling and 3D printing with the goal of understanding and supporting students' spatial thinking during computer-aided designs. As students solved an authentic engineering design problem that required making 3D models of prosthetics for animals with disabilities, they developed technological literacy in computer-aided modeling tools as well as 3D printing technology. Meanwhile, students exercised their spatial skills such as mental rotation, perspective-changing, and forming mental models of their digital designs.

A related line of work is illustrated by Peng and Sollervall (2014), who developed a technology-supported, outdoor learning activity that engaged students in learning mathematics concepts, solving a real-world mathematical problem, and exercising their spatial orientation skills. Using a mobile application that informed students of the relative distance between themselves and the physical markers on the field, sixth-grade students navigated the outdoor environment, continuously estimated and calculated distances, and tested out different spatial orientation and coordination strategies to solve the math tasks.

1.2.4 Engaging students in spatial thinking through engineering design

Engineering design projects offer a promising platform to integrate knowledge needed for different disciplines while challenging students' inquiry, analytical, and problem-solving skills (Kelley & Knowles, 2016). Several interventions that target students' spatial ability development have been delivered in the form of engineering design projects (e.g., Goktepe Yildiz & Ozdemir, 2020; Ramey & Uttal, 2017; Taylor & Hutton, 2013). For example, a robotics programming intervention for fourth- to sixth-grade pupils led to improvements in a range of spatial reasoning skills, such as building 3D objects from pictures and shape rotation (Francis et al., 2021). These gains in the spatial task performance were observed in both the short-term, week-long group and the long-term, year-long group.

Similarly, Julià and Antolì (2018) developed a multidisciplinary STEM course on robotics to develop students' spatial skills and mechanical reasoning skills. This year-long course for sixth- and seventh-grade pupils used educational robotics kits to promote hands-on learning, through which pupils worked in groups to solve mechanical modeling and construction problems. This robotics course led to statistically significant increases in perspective-taking spatial orientation task performance of these pupils. Classroom observations further indicated that the course supported pupils' learning of not only engineering and technology concepts, but also those related to science and mathematics. Despite the small sample size, these findings demonstrate that a STEM course that is integrative in its nature has the potential to develop both spatial skills and STEM knowledge among students.

Within integrated STEM education, engineering and technological design practices often serve as a core through which math and science concepts are connected (National Academy of Engineering & National Research Council, 2014; Kelley & Knowles, 2016; Sanders, 2009). Empirical research has provided evidence on implementing design-based, integrated STEM projects to enhance students' knowledge in science (Wendell & Rogers, 2013), mathematics (Burghardt et al., 2010), STEM concepts (Fan & Yu, 2017), and spatial skills (English, 2019). Taken together, these studies indicate that design-based approaches function as an important bridge in integrated STEM education, providing a practical pathway for supporting meaningful integration across scientific knowledge, technology literacy, engineering practice, and mathematical reasoning (de Vries, 2021; Cook & Bush, 2018; Crismond & Adams, 2012; Hallström & Ankiewicz, 2023). Nevertheless, despite design's critical role in integrated STEM, its relationship with spatial thinking remains insufficiently understood—a gap this dissertation seeks to address.

1.3 Design and spatial thinking: what is known and what remains to be understood?

The capacity to turn our creative visions into designs and technologies has been a transformative force throughout history (Klapwijk & Stables, 2023). Essentially, design means conceptualizing and actualizing new entities (Cross, 2006). It involves processes such as defining and redefining problems, understanding users' needs, making prototypes through various media, and developing, testing, and iterating on prototypes. Numerous studies have underscored the importance of design in general education (Cropley & Cropley, 2009; Cross, 2006; Kimbell & Stabes, 2007), and its growing prominence in primary and secondary curricula worldwide is evident in regions such as Europe (Kimbell & Stabes, 2007), North America (Goldman et al., 2009), and Asia (Koh et al., 2015).

In the dynamic and iterative process of designing—where the minds and the hands interact together to shape ideas and products—designers make use of tools and skills, such as problem framing (Schön, 1983), context mapping for user research (Visser et al., 2005), analogical thinking (Hey et al., 2008), sketching (Purcell & Gero, 1998), as well as spatial thinking (Ramey & Uttal, 2017; Suh & Cho, 2020). These practices support the creation of designs that are innovative, relevant, and meaningful to industries and societies. In the historical analysis of technological developments, Ferguson (1977) proposed that 'thinking with pictures' and imagining the assembly and alterations of objects and elements in the mind are inherent in engineering and mechanical art practices. Historically, scientists, engineers, designers, and technologists frequently leveraged these types of nonverbal thoughts. As Cross (2006) outlined, one of the core design abilities involves using 'nonverbal, graphic/spatial modelling media' (p. 20), demonstrating the longstanding role of visual-spatial thinking in design practice and design education.

1.3.1 How expert and novice designers use spatial thinking in design

An emerging body of design research has provided qualitative or mixed-methods evidence on how designers use their spatial thinking in the process of designing, especially during problem exploration, idea emergence, and sketching, where attention to visual details supports iteration and refinement. Cross (2006), for example, documented the thought processes of expert designers in design problem solving. In one case, an experienced designer was asked to design 'a carrying/fastening device that would enable you to fasten and carry a backpack on a mountain bicycle' (p. 64). This designer began with exploring and analyzing the design problem by visualizing the problem context in mind. By imagining the different movements performed by the rider going uphill and downhill, and the possible positions where the backpack could be attached in this dynamic process, the designer visualized from different perspectives and examined the spatial positions and relations between the rider, the bike, and the surrounding environment. Following problem exploration, the emergence of a new design idea, such as perceiving an ambiguous form of a new product in one's mind and continuing to conceptualize it into a more concrete mental image, can be the result of a series of interactions between perception, visual memory, mental imagery, and mental transformation (Oxman, 2002). To externalize mental images of ideas, designers typically produce sketches of preliminary ideas and concepts. This practice allows expert designers to attend to the visual details of their ideas, through which they can discover and utilize the spatial and functional relations entailed, thereby generating new ideas or reconfiguring existing visual-spatial elements (Suwa & Tversky, 1997).

Novice designers, such as K-12 pupils, also utilize spatial thinking in design processes. In architecture design projects, such as planning cities and building playgrounds, pupils are often immersed in the experiences of different forms of spaces, which require them to carefully consider spatial relations to satisfy the needs of different groups of people (Acer, 2016; Batic, 2011; Seitamaa-Hakkarainen et al., 2012). Other design and maker projects may have a stronger focus on the use of technology, such as primary school pupils engaging in virtual 3D makerspaces (Fowler et al., 2024) or secondary school pupils designing through 3D modeling tools (Bhaduri et al., 2019; Dilling & Vogler, 2021). These contexts challenge pupils' ability to visualize and create 3D representations of their ideas using computer-aided design tools. Design projects that emphasize hands-on tinkering and iteration, including secondary school pupils designing, building, and testing functional windmills (Khunyakari et al., 2007) and preschoolers engaging in playful making with materials to solve an open-ended, story-based design problem (Sonneveld et al., 2024), also recruit various practices of spatial thinking, such as visualization, 2D-3D spatial transformation, and spatial motor movements. In addition, in engineering design activities, secondary school pupils have been shown to use not only spatial thinking to visualize creative solutions to engineering problems, but also language, sketches, gestures, and hands-on interactions with objects to facilitate their spatial understanding and enrich their visuo-spatial representations (Ramey & Uttal, 2017).

1.3.2 Addressing gaps in spatial ability research through the lens of design education

Recent research has increasingly addressed the important role spatial thinking plays in learning STEM (Buckley et al., 2018; Newcombe, 2017; Wai et al., 2009). By contrast, considerably less is known about the role of spatial thinking in design practices, even though these practices are among the key tools used in STEM problem-solving. This dissertation sets out to address this research gap by (1) exploring how K-12 pupils use spatial thinking in design projects and (2) examining the relationship between pupils' spatial ability and their design performance. Clarifying these relationships will provide insights into how design-based thinking and practices can be leveraged to cultivate pupils' spatial skills and, in turn, enhance their engagement and achievement in STEM.

Many of the studies reviewed above have situated students' spatial skills development in authentic STEM problems and reported both quantitative data, such as spatial test scores, and qualitative data drawn from classroom observations and interviews. Combining these methods brings forward insights that cannot be captured through quantitative measures alone, shedding light on not just what was learned but also how learning unfolded. This dissertation will build on these insights by using qualitative findings to guide subsequent mixed-methods and quantitative study design and analysis. By synthesizing insights from qualitative, quantitative, and mixed-methods studies, this dissertation aims to provide a nuanced understanding of how pupils employ spatial thinking in design education contexts and of the relationship between spatial ability and design performance.

1.4 Research outline

To answer the main research questions of how pupils use spatial thinking in design and how spatial ability influences design quality, this dissertation comprises two qualitative case studies, one mixed-methods study, and one quantitative study.

Chapter 2 addresses the research question: How do pupils use spatial thinking to create tangible visualizations in design and making? In today's data-driven world, understanding and effectively communicating data is an important educational topic for young learners (Ben-Zvi & Garfield, 2004). Data physicalization making, which involves encoding intangible data variables with visual, spatial, and physical properties, provides an engaging, accessible, and hands-on approach that inherently draws on spatial thinking (Huron, Jansen & Carpendale, 2014; Vande Moere & Patel, 2010). While past studies have investigated learning from and with infographics—such as how students' ability to interpret two-dimensional diagrams relates to their spatial ability (Kragten et al., 2015)—little is known about how pupils apply spatial thinking to create three-dimensional, physical data visualizations. This case study examines how upper-primary school pupils created physical data visualizations that embodied their understanding of both spatial and numerical relationships.

Chapter 3 focuses on a Design and Technology (D&T) project and addresses the research question: How do pupils use spatial thinking in an open-ended, analogy-driven design process? Much research has linked spatial ability to STEM learning (e.g., Newcombe, 2017; Uttal & Cohen, 2012), primarily emphasizing the use of spatial thinking in solving analytical, often closed-ended, problems. Design, on the other hand, tackles problems that are ill-defined and open-ended (Cross, 2006). Instead of identifying a single correct answer, designers may use spatial thinking in a distinct way to envision and bring to life products that do not yet exist. Studies on designers' cognition (e.g., Schön & Wiggins, 1992; Suwa & Tversky, 1997) have illustrated various types of visual-spatial thinking involved in the design process; however, less is known about how novice designers, such as young pupils, use spatial thinking in design. Moreover, the use of analogy as both a spatial reasoning tool (Gentner et al., 2016) and a design ideation tool (Casakin & Goldschmidt, 1999) remains underexplored in authentic classroom contexts. This chapter adopted a design-based research approach to develop a biomimicry design project that engaged pupils in spatial thinking. Classroom observations of the project provided qualitative insights into how lower-secondary school pupils' spatial thinking unfolds during an open-ended, design-by-analogy task.

Insights from these two case studies—focused on STEM-related making and open-ended design problem-solving—together with emerging literature in the field, suggest that spatial thinking plays a meaningful role in design. To extend these findings and examine how spatial ability influences creative design outcomes, a quantitative study was conducted. While prior research has investigated how pupils develop spatial ability through computer-aided design (Fowler et al., 2024) and engineering activities (Goktepe Yildiz & Ozdemir, 2020), relatively little is known about how spatial ability influences the early stages of creative design, particularly ideation and prototyping. Accordingly, the third and fourth studies examined how spatial ability and creativity contribute to the generation and development of creative design ideas.

Chapter 4 addresses the methodological challenge of evaluating creativity in design work through the research question: How do expert judges evaluate creativity in pupils' design ideas and prototypes through comparative judgement, and how do their evaluative criteria evolve over time? Although D&T classrooms offer rich opportunities for creative expression, adequately assessing creativity is often complicated by the open-ended nature of design tasks and the varied procedures involved in design (Kimbell & Stables, 2007). Comparative judgement as an assessment method is considered well-suited for assessing complex, hard-to-define constructs like creativity (Jones & Alcock, 2014). Different from scoring with predefined rubrics, comparative judgement relies on experts using their professional judgement to holistically compare and identify the relative quality of pupils' work (Hartell & Buckley, 2021). This chapter combines qualitative and quantitative analyses to offer an alternative lens on the multifaceted nature of design creativity.

Building on the assessment results, Chapter 5 seeks to answer the research question: Does spatial ability relate to pupils' creative design performance, and is this relationship direct or mediated through creativity? Existing research has reported mixed findings regarding the relationship between spatial ability and creative design performance among university students (Cho, 2017; Reid & Sorby, 2023; Suh & Cho, 2020), while its influence on younger learners remains understudied (Chang, 2014). As design is gaining recognition as a discipline that brings forward creativity and innovation (Cropley & Cropley, 2010), and as design-based practices are becoming increasingly prevalent in STEM problem solving (Cook & Bush, 2018; Hallström & Ankiewicz, 2023), there is a need to clarify the respective contributions of spatial ability and creativity to pupils' creative design performance.

The final chapter synthesizes findings from the four empirical studies to advance current understanding of the relationship between spatial thinking and design education. By examining both hands-on design activities and cognitive predictors of design performance, this dissertation highlights the diverse ways through which spatial thinking is recruited in authentic, open-ended tasks. Focusing on younger learners and the early phases of design—areas that remain underexplored in existing research—this work offers empirical insights and identifies promising directions for future research and educational practice.

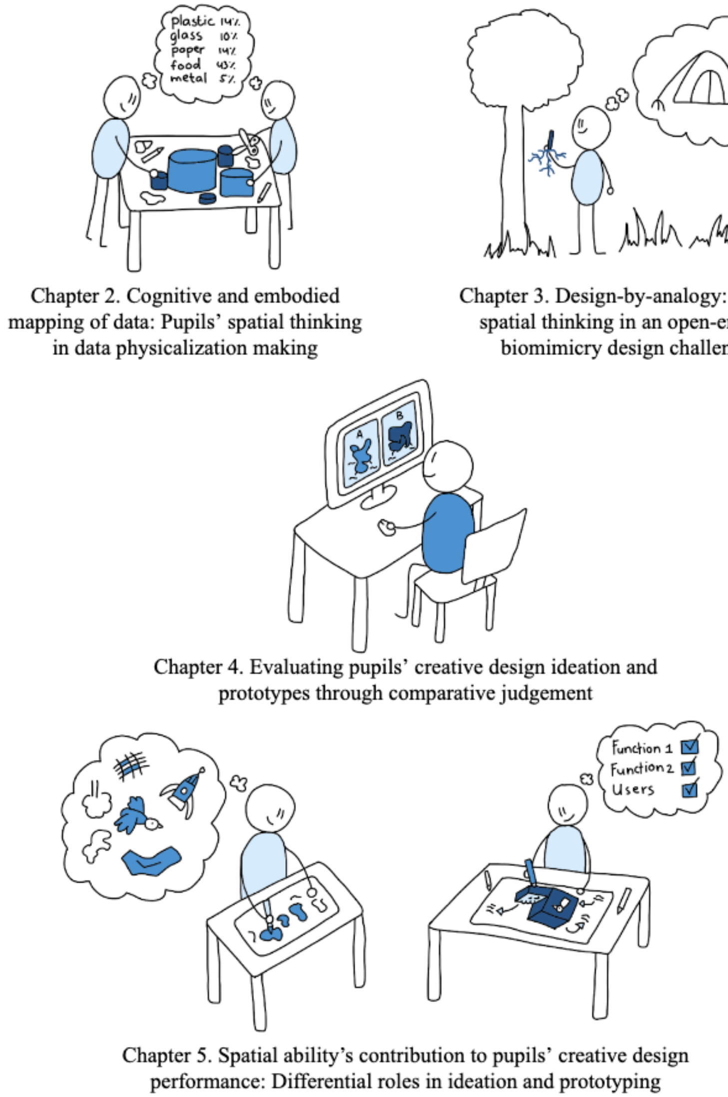


Figure 1.1 Visual overview of four empirical studies in this dissertation

CHAPTER 2
Cognitive and embodied mapping of data: Pupils' spatial thinking in data physicalization making

CHAPTER 2
Cognitive and embodied mapping of data: Pupils' spatial thinking in data physicalization making[2]

Understanding and effectively using visual representations is important to learning science, technology, engineering, and mathematics. Various techniques to visualize information, such as two- and three-dimensional graphs and diagrams, not only expand our capacity to work with different types of information but also actively recruit our visual–spatial thinking. Data physicalization is emerging as a beginner-friendly approach to constructing information visualization. Mapping intangible data onto tangible artifacts that possess visual, spatial, and physical properties demands an interplay of spatial thinking and hands-on manipulation. Much existing literature has explored using formatted infographics to aid learning and spatial thinking development. However, there is limited insight into how school-age pupils may leverage their spatial thinking to create information visualizations, particularly tangible ones. This case study documented the data physicalization activities organized in two design classrooms of an international school in the Netherlands, with 37 pupils aged 11–12. Seven themes relevant to spatial thinking were identified from multimodal evidence gathered from the data physicalization artifacts, classroom videos, and recordings of pupils' making process, and semi-structured interviews with pupils. Findings from this study suggested that these pupils generated various ideas to create visual–spatial forms for data with the materials at hand, such as mapping quantities to tangible materials of different sizes, using spatial ordinal arrangement, and unitizing materials to set visual parameters. Meanwhile, they evaluated and adjusted the visual–spatial properties of these materials according to the numerical data they had, crafting feasibility, and others' spatial perspectives. What was particularly interesting in our findings was pupils' iteration on their visual–spatial understandings of the intangible numerical values and the tangible materials throughout the embodied making processes. Overall, this study illustrates the different types of spatial thinking pupils applied to create their data physicalizations and offers insights into how embodied experiences accompanying the open-ended visualization challenge allowed pupils to explore and construct spatial understandings.

---

[2] *This chapter includes text that has previously appeared in the following source:*

Zhu, C., Klapwijk, R., Silva-Ordaz, M., Spandaw, J., de Vries, M. J. (2023). Cognitive and embodied mapping of data: an examination of children's spatial thinking in data physicalization. *Frontiers in Education, 8,* 1308117.

2.1 Introduction

Growing up in a world where information is ubiquitous, children navigate through diverse visual representations of data in their daily encounters, whether it is checking nutrition labels on cereal boxes, deciphering scavenger hunt maps, or managing inventories in video games. The need to comprehend and effectively communicate information, such as data, is not only relevant but also important to children's educational and social development (Ben-Zvi & Garfield, 2004; Dasgupta & Hill, 2017). A common resort is to use visual or tangible representations, which allow us to externalize products of thinking, aid our understanding of knowledge (Tversky, 2001; Ainsworth, 2006), and pave the way for subsequent inspection and manipulation (Fish & Scrivener, 1990; Kirsh, 2009).

Graphs and diagrams are some of the common means of visualization that permeate STEM disciplines. Importantly, the learning of these visualization techniques often requires spatial thinking (Tversky, 2001; Kim & Maher, 2008; Hegarty, 2010; Höffler, 2010; Newcombe, 2010, 2017; Stieff et al., 2010; Frick & Newcombe, 2015; Baykal et al., 2018). Spatial thinking can be defined as "the mental process of representing, analyzing, and drawing inferences from spatial relations between objects or within objects" (Uttal, Meadow, et al., 2013, p. 367) and is widely considered important to the learning of science, technology, engineering, arts, and mathematics (Sorby, 2009; Wai et al., 2009; Uttal, Miller & Newcombe, 2013; Buckley et al., 2018; Zhu, Leung et al., 2023). Childhood and early adolescence mark an important period of rapid spatial ability development (Thurstone, 1955; Newcombe & Frick, 2010; Maresch & Sorby, 2021). Several studies have explored children's understanding and use of visual–spatial representations in relation to their spatial skills (Szechter & Liben, 2004; Hannafin et al., 2008; Frick & Newcombe, 2015). Only a select few studies shed light on how children's creation of visualizations reflected their spatial thinking (Diezmann & Watters, 2000; Van Meter et al., 2006; Tytler & Prain, 2010), with pupils in grades one through six, approximately aged 6 to 12. Despite these studies, there remains a need for more research to understand how school-age pupils learn, interact, and construct visual–spatial representations, as well as the potential roles spatial thinking plays in these types of learning (Ramadas, 2009; Bergey et al., 2015).

Following the call to make information visualization accessible to novices (Huron, Carpendale, et al., 2014), one type of information visualization technique is gaining increasing attention from researchers, data artists, and educators because of its value in education (Vande Moere & Patel, 2010) and public communication (Taylor et al., 2015). Data physicalization is defined as "a physical artifact whose geometry or material properties encode data" (Jansen et al., 2015, p. 3228). Unlike paper-based or screen-based visual representations, data physicalization leverages physicality and represents abstract information in data variables through properties of tangible objects, such as their forms, dimensions, and positions (Zhao & Vande Moere, 2008). Data physicalization activities are generally engaging and stimulate creativity, and have the potential to promote understanding of data even among novices like children (Bhargava & D'Ignazio, 2017; Kanis, 2019; Bae et al., 2022). Crafting or engaging with data physicalizations taps into our visualization and perceptual exploration competencies (Baker et al., 2009; Huron, Carpendale et al., 2014; Huron, Jansen & Carpendale, 2014; Jansen et al., 2015; Jansen & Hornbæk, 2015; Wun et al., 2016; Hull & Willett, 2017). In addition to providing visual cues, data physicalization affords tangible interactions that leverage our active perception skills (Jansen et al., 2015). For example, we can move our bodies to inspect physicalizations from different angles and distances. By organizing visual and tangible cues in a spatial manner, data physicalization creators make use of the spatial relations between data variables (Kirsh, 2009; Hull & Willett, 2017) to highlight meaningful associations (Lin, 1992). On the other hand, viewers apply their spatial perception skills to decipher details from visual and tangible indicators such as shape, volume, and spatial position (Jansen et al., 2015).

This case study aimed to understand how pupils aged 11–12 make sense of data through their spatial thinking and delved into the diverse ways pupils utilize spatial thinking during their data physicalization construction (see Figure 2.1 for an example). Specifically, given that the use of tangible materials distinguishes data physicalization from many conventional paper-based or screen-based visualizations, we also investigated how embodied interactions with tangible materials of visual–spatial properties could engage and support pupils' spatial thinking.

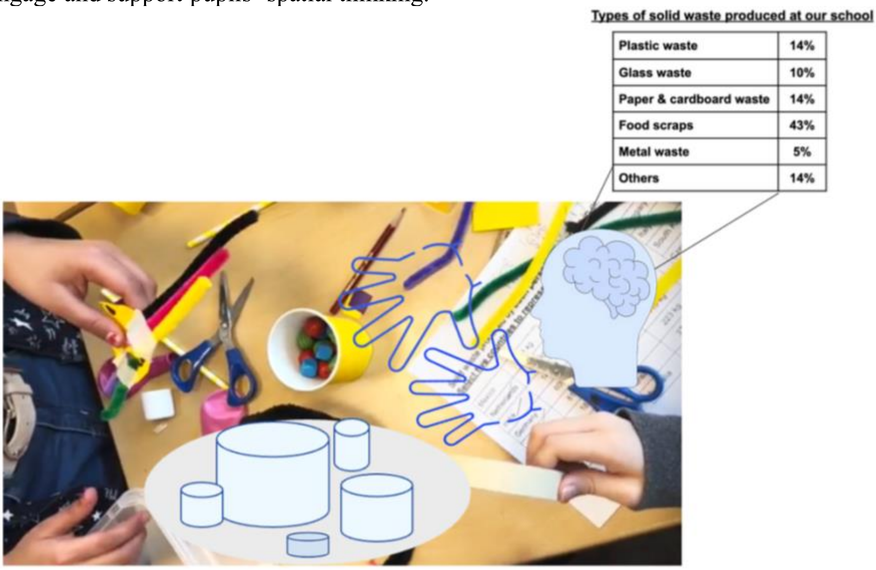


Figure 2.1 An illustration of pupils' construction of data physicalization

2.2 Literature review

2.2.1 Visual–spatial representations permeating STEM disciplines

Visualizations such as cladograms in biology, ball-and-stick models in chemistry, and pulley diagrams in physics are essential tools for grasping abstract STEM concepts. Drawing on Mayer's (2005) cognitive theory of multimedia learning, processing visualizations employs a range of our cognitive functions, such as visual perception and the visuospatial sketchpad, to select, organize, and integrate verbal or textual information into visual representations in the mind, and to establish the referential mapping between them. Spatial thinking appears to be an important factor in ensuring that students effectively utilize these visualization tools to traverse different disciplines (National Research Council, 2006).

The ability to recognize, comprehend, and make use of the information denoted by visualizations develops from a young age (Frick & Newcombe, 2015). Existing literature has mainly explored the relationship between K-12 students' use of formatted, two-dimensional screen-based or paper-based infographics and their spatial thinking (e.g., Hannafin et al., 2008; Cromley et al., 2013; Frick & Newcombe, 2015; Kragten et al., 2015). Several studies have indicated that secondary school students' (7th to 12th grades) self-generated visual–spatial representations are helpful tools to facilitate their science learning (e.g., Stieff, 2011; Leopold & Leutner, 2012; Cooper et al., 2017; Tytler et al., 2020). Concurrently, a selection of studies has also shed light on how self-generated visual–spatial representations recruited primary school students' spatial thinking and helped them make sense of abstract concepts. For instance, when sixth-grade students were challenged (with support) to develop visual representations based on verbal descriptions of a bird's wing feather placement, they utilized the spatial information provided and spatially located layers of feathers on different wing sections, consequently showing increased knowledge about the structure and mechanics of wings (Van Meter et al., 2006). A longitudinal study by Tytler and Prain (2010) reported that students progressively iterated on their spatial representations of water evaporation from grade one to six, as their understanding of this phenomenon deepened. Moreover, through a series of lessons emphasizing the use of representations, fifth and sixth graders leveraged different modes of visual–spatial representations, such as sketches, graphs, animations, and 3D models, to document their understanding of the dynamic process of water evaporation at the molecular level, demonstrating enriched science reasoning and active use of spatial reasoning (Tytler et al., 2013).

Recently, a meta-analysis conducted by Hawes et al. (2022) examined the impact of spatial interventions, discovering that spatial training incorporating hands-on manipulatives was more effective in enhancing math performance compared to training delivered in digital or paper-pencil formats without the use of concrete materials. This finding is especially intriguing, considering a primary objective of spatial training is to enhance performance in STEM fields (Newcombe, 2010; Uttal, Miller & Newcombe, 2013; Buckley et al., 2018; Zhu, Leung et al., 2023). It raises questions about the role of hands-on materials in fostering spatial thinking development, especially among young students. In the following section, we will review how hands-on experiences support spatial thinking and explore the prospect of understanding primary school students' spatial thinking through data physicalization—a visualization technique that draws from direct physical manipulation and embodied interactions.

2.2.2 Spatial understandings through embodied cognition

A primary means by which children make sense of and learn about their environment is by interacting with physical objects (Piaget, 1952; Inhelder & Piaget, 1964). The embodied cognition perspective, suggesting that our cognitive activities are anchored by our sensorimotor interactions with the surroundings (Wilson, 2002; Noë, 2004; Barsalou, 2008), has gained attention in various disciplines. Friedrich (1782–1852) was one of the first to take the initiative of integrating hands-on materials into the classroom to encourage self-directed playful learning among young children (Liebschner, 2006). What was later known as Fröbel's gifts consisted of materials such as wooden spheres, cubes, planks, and prisms of different sizes and quantities. These materials allow for various possible ways of manipulation and spark children's exploration of the properties of different shapes, dimensions, and balance. More recently, research has explored how hands-on, bodily experiences facilitate the learning and externalization of abstract knowledge (O'Malley & Fraser, 2004; Marshall, 2007; Shaer & Hornecker, 2010), such as in physics (Kontra et al., 2015), mathematics (Tsang et al., 2015), and chemistry (Stull et al., 2018). In the field of design and making, particularly, the hand plays a critical role as it "grasps the physicality and materiality of thought" (Pallasmaa, 2017, p. 104). Constructing and interacting with three-dimensional prototypes allows designers—from children to adults—to gain spatial insights and iterate on design concepts and forms (Groth & Mäkelä, 2016).

Spatial thinking, which involves many highly abstract processes that occur in the mind's eye, also relies on spatial motor movements as one of its basic practices (Maresch & Sorby, 2021). These movements often represent an extension of the visual–spatial perception and cognitive reasoning happening in the mind. Before children develop the capacity to form abstract, verbal representations, their internal representations in the early years, as Kosslyn (1978) theorized, predominantly consist of information represented through actions and mental imagery. Hands-on manipulations, for example, can elicit motor and visual feedback on our mental imagery, enabling us to assess and modify our perceptual and spatial understanding of objects or situations (Frick et al., 2009). Through a series of experiments using the task of visualizing the water levels in tilted glasses, Frick et al. (2009) found that manually tilting the glass, as opposed to imagining the tilting in mind, supported the mental imagery of physical relations and the mental transformation of water levels among children as young as 5 years old. Link et al. (2013) discovered that having first-graders walk along a physical number line and estimate the spatial position of numbers enhanced their numerical competencies. A number of other spatial interventions have leveraged tangible resources such as paper-folding, block-building, and educational robotics to foster pupils' development of spatial thinking (Casey et al., 2008; Burte et al., 2017; Julià & Antolì, 2018). For example, Burte et al. (2017) developed a series of paper engineering activities that engaged third-to-sixth-grade pupils in folding and cutting papers to create three-dimensional pop-ups. These embodied activities led to enhancements in pupils' spatial and math task performance, especially for those in higher grades. There are many explanations for why embodied elements can support spatial thinking. For instance, hands-on manipulation can aid problem-solving when mental manipulation imposes too much cognitive load (Wilson, 2002; Shaer & Hornecker, 2010; Stull et al., 2018). Besides, tangible materials can cue mental images that enrich and refine the scope of problem exploration and solution searches (Stigler, 1984; Chao et al., 2000). There remain concerns that hands-on experiences might result in learning outcomes that are primarily due to the scaffolding provided by the manipulatives (Newcombe, 2017). Nevertheless, as the above-mentioned studies show, embodiment appears to play a meaningful role in the spatial development of pupils in different grades.

To date, the benefits and challenges of using visual displays, such as two- and three-dimensional graphs, diagrams, and models, to facilitate students' subject learning and spatial skills development have been explored widely (e.g., Wu & Shah, 2004; Stieff et al., 2005). However, the understanding of school-age pupils' interaction with tangible visualization—which incorporates both visual and hands-on elements—is still limited and warrants further investigation.

2.2.3 Data physicalization and constructionism

Many existing spatial interventions give children tasks that are analytical in nature, with closed-ended, pre-determined answers. Yet from the constructionism viewpoint, children spontaneously develop knowledge and comprehension by engaging with and creating things from their physical and social surroundings (Papert, 1991, 1994), a process that is typically hands-on. Open-ended activities such as data physicalization therefore provide a window into how children construct their understanding of data by creatively using materials that have diverse visual–spatial properties and can be arranged to represent various spatial relations.

Data physicalization is essentially a process of constructing understandings of intangible data through tangible materials. It is inherently an embodied process, echoing what Wilson (2002) termed as off-loading mental activities onto the materials and the surrounding space. Unlike iconic visual–spatial displays, such as a map, where things being represented already have visible spatial properties, data physicalization often requires the maker to encode numerical information that is not inherently spatial in visual–spatial ways, and requires the viewers to decode the non-spatial information from a spatial representation of number magnitude. Drawing from the information visualization process model (Card & Mackinlay, 1997; Jansen & Dragicevic, 2013), transforming abstract data into visual or tangible representations requires visual mapping. This means assigning visual properties—which can embody spatial or graphical properties (Card & Mackinlay, 1997)—to data variables. Then, the maker carries out presentation mapping to refine the data display, such as tinkering with the way of presentation to facilitate viewers' perception and understanding (Jansen & Dragicevic, 2013). Although this model does not depict all cognitive processes involved in constructing data physicalization, it shows that mapping intangible information and relations onto visual–spatial forms inherently requires spatial thinking, which, as discussed earlier, involves understanding, representing, and making meanings out of the spatial relations between and within objects.

Several studies have explored how visual–spatial thinking is used when constructing or interacting with data physicalization. For instance, variances in numerical data can be mapped to physical qualities such as sizes (e.g., a line of a certain length, or a 3D shape of a certain volume) (Vande Moere & Patel, 2010; Jansen & Hornbæk, 2015). Makers of data physicalization may determine the layout of the physicalization by allocating unit values to specific dimensions (e.g., lengths) or quantities of tangible objects. These visual parameters can aid in both their processing and depiction of spatial relations among data variables (Huron, Carpendale et al., 2014; Huron, Jansen & Carpendale, 2014). In addition, spatial arrangement and ordering can highlight hierarchical or ordinal relationships in data variables (Vande Moere & Patel, 2010; Huron, Jansen & Carpendale, 2014), reflect an efficient spatial layout (Perin, 2021), or assist viewers in comprehending and organizing information (Wun et al., 2016). Interestingly, Huron, Jansen, and Carpendale (2014) suggested that arranging tangible materials might impact people's internal spatial organization, namely their mental representation of the spatial relationship between data points and physical objects. Overall, compared to paper-based or screen-based information visualizations, data physicalization offers tangible and embodied interactions that enrich perceptual insights (Zhao & Vande Moere, 2008; Jansen et al., 2013, 2015; Search, 2015) and prompt an inspection into the spatial organization of visual and tangible elements (Wun et al., 2016). The resulting physicalization artifacts in turn embody such spatial understandings. Notably, the low-tech, tangible resources employed in data physicalization are especially suitable for empowering young novices like children to experiment with diverse concepts and make modifications handily (Huron, Jansen & Carpendale, 2014; Bhargava & D'Ignazio, 2017; Bae et al., 2022).

Numerous data physicalization workshops have been designed for students in higher education (Vande Moere & Patel, 2010; Huron et al., 2016; Willett & Huron, 2016; Hurtienne & Reinhardt, 2017; Perin, 2021). Yet only a limited number of studies (e.g., Bhargava & D'Ignazio, 2017; Kanis, 2019; Bae et al., 2022) have delved into how younger children craft data physicalizations. These studies have primarily focused on children's and pupils' creativity, engagement, and understanding of data during the physicalization process, rather than examining their spatial thinking. Through this case study, we aimed to contribute to the ongoing discussion by leveraging the embodied learning experiences in data physicalization to engage school-age pupils in spatial thinking. The research questions to be addressed through our qualitative examination are (1) How does the process of pupils mapping numerical data onto tangible materials reflect their spatial thinking, and (2) How does the embodied process of making data physicalization stimulate pupils' spatial thinking?

2.3 Methods

2.3.1 Participants

This study involved 37 pupils (18 girls, 19 boys), aged 11 to 12, from the final year of primary school at an international school in the Netherlands. We selected these pupils because their Design and Technology teacher collaborated with the researchers to integrate data physicalization activities into the classroom. All participants provided informed consent to be part of this study. By this grade, pupils have knowledge of multi-digit numbers, multi-step addition and subtraction, multiplication and division, fractions and percentages, characteristics of 2D and 3D shapes, units of measurement and simple unit conversion, collecting, displaying, and interpreting data in data tables, line graphs, bar graphs, and pie charts (International Baccalaureate, 2009), which may be relevant to data physicalization construction. This study targeted students at the primary school level for several reasons. Firstly, it is important for younger students to begin developing visual literacy (Prain & Waldrip, 2010; Tytler & Prain, 2016) as well as statistical and data literacy (Ben-Zvi & Garfield, 2004; Mickelson & Heaton, 2004). There remains a need to better understand how primary school students fabricate visual and tangible representations from data. Moreover, primary education places a strong emphasis on hands-on learning. Studies have shown that pupils from third to sixth grades benefited from an embodied spatial training, with those at the upper primary level exhibiting more improvements in their math performance than younger pupils (Burte et al., 2017). Lastly, the primary curriculum is more flexible than secondary education, making it an ideal platform for integrating data physicalization activities into the classroom.

2.3.2 Materials

In the data physicalization activity, pupils used two primary categories of materials for construction: data tables and crafting materials. The data tables, curated to align with the design theme, Designing for Circular Economy, contained only the necessary information to highlight the goal of the task (Waldschütz & Hornecker, 2020). Data values were presented in non-ordered form with units of measurement suitable for this age group, such as kilograms and percentages. The pupils had the option to select from data tables detailing: (1)The percentages (by weight) of five different waste categories produced at their school in 2020; (2) The percentages (by weight) of five different waste categories generated in their native countries, sourced from the World Bank database (2018); (3) The amount (by weight) of solid waste produced per capita in 23 different countries, also sourced from the World Bank database (2018); choose five countries to represent.

A variety of crafting materials, including playdough, elastic bands, straws, pipe cleaners, paper cups and plates, yarn, beads, and balloons, were provided to enable a range of possible operations (Huron et al., 2017), such as molding, bending, tying, and stacking. These materials were familiar to the pupils, which should help lower the barrier to starting the activity (Bhargava & D'Ignazio, 2017).

2.3.3 Procedure

The data physicalization activity was integrated into the Designing for Circular Economy curriculum unit, a ten-session design project created by World's Largest Lesson in partnership with UNICEF, and made use of the Your Turn to Make Your Mark in Design approach developed at TU Delft. In this design project, pupils explore and analyze the problem of waste, and develop design ideas by reusing, repairing, or repurposing an object for users with different needs in their family, school, or community. They then make, test, and present their designed prototypes. Collaborating with the Design and Technology teacher, the research team integrated the data physicalization activity into the first two sessions of the design project, aiming to have pupils explore waste-related data and gain a deeper understanding of the importance of designing through reusing and repurposing waste. The activities were organized in two classes: the first class of 18 pupils in the morning and the second class of 19 pupils in the afternoon, both led by the same Design and Technology teacher. The pupils were instructed to make creative and easy-to-understand data physicalizations based on their selected data tables. Pupils were free to choose to work in pairs or individually, as they usually do in their design units. On average, pupils from both classes had 15 minutes to read data tables, develop initial ideas, and select crafting materials. They then dedicated around 40 minutes to making the tangible, visual representations of data. The activity concluded with a 15-minute session where pupils presented their data physicalization artifacts to the class. Due to time constraints in the school schedule, only pupils in the afternoon class had time for in-class presentations. During the class, the researchers occasionally asked open-ended questions to pupils for clarification on their ideas or their making process.

2.3.4 Data collection and analysis

A total of 23 data physicalization artifacts were created by the pupils: 14 were made in pairs and 9 were made individually. Upon finishing their data physicalizations, pupils filled out a five-question self-report survey to reflect on their experience in the data physicalization activity. They gave Likert-scale ratings to five statements regarding the process of their data physicalization construction and their understanding of data. The statements were as follows: "Before making, I thought about how the visualization would look"; "During making, I checked to see if the visualization represents the data well"; "I can explain the meaning of the data to others using the visualization"; "I can understand the data better now"; "I have more ideas for what I want to design for the design challenge." 22 self-report surveys were collected. Of all the pupils, 22 who consented to interviews participated in brief semi-structured interviews. They were asked open-ended questions about how they developed ideas and visualized their physicalizations in mind, how they made key decisions about and changes to their physicalizations, how they thought their physicalizations could help others understand the data, and any challenges they faced.

A total of 6 hours and 18 minutes of video/audio recordings were collected from the classroom and the semi-structured interviews. Transcriptions of classroom and interview recordings were divided into segments, with each segment capturing pupils' discussion on a specific topic before transitioning to the next. These conversation segments were analyzed jointly with pupils' data physicalization artifacts, constituting episodes of data. Four researchers contributed to the iterative process of qualitative coding. The initial round of data-driven thematic analysis was carried out on MAXQDA by the principal researcher, rendering a preliminary list of themes. In the second round of coding, three other researchers were invited to independently evaluate the congruence of the themes with randomly selected episodes of data. This included a researcher in design education who is involved in the study, a researcher in mathematics education with less involvement, and another researcher in spatial language from a different institution with no involvement in the study. The four researchers discussed any disagreements regarding themes and coding interpretations. Using the iterated themes, the four researchers carried out the third round of coding, where each of them independently coded seven randomly selected data episodes. Fleiss kappa, which measures the proportion of observed agreement over and above agreement expected by chance (Fleiss & Cohen, 1973), was calculated to determine the inter-rater reliability between the four coders. The kappa value was 0.728, 95% CI [0.613, 0.842] with a significance level of $p < .001$, indicating a high level of agreement (Landis & Koch, 1977). Any remaining differences in coding were then discussed among researchers, culminating in agreement on the final list of themes as well as interpretations of the randomly selected episodes. Through the data-driven coding process, seven themes were identified to depict the different types of spatial thinking reflected in pupils' data physicalization artifacts—Spatial Ways of Representation, and through their making processes—Constructing Spatial Understanding through Embodied Making.

2.4 Results

All pupils were able to create data physicalizations either in pairs or individually. All but two of the data physicalization artefacts explicitly demonstrated topological relations among the data variables and could align with one or more of the spatial thinking themes we identified, either in their ways of representation, their making process, or both. The episodes presented in Table 2.1 and in subsequent paragraphs came from both in-class observations and semi-structured interviews with pupils. Note that all pupils' names mentioned in these episodes have been pseudonymized. While each episode may align with more than one theme, it is listed under the theme that it reflects most prominently.

Table 2.1 Pupils' demonstration of spatial thinking in their physicalization artifacts and during their making processes

Theme	Definition	Example
Spatial Ways of Representation		
Quantities-to-sizes mapping	Pupils representing value differences in the numerical data they selected by varying the sizes (e.g., length, width, height, volume) of tangible materials through measuring or estimation	"And like, I did the biggest one for the biggest amount. And then I see, like the rest of the pom poms, how much smaller they need to be." —Nia
Spatial ordinal arrangement	Pupils highlighting ordinal relationships in the numerical data they selected by	"Maybe from smallest to biggest, instead of biggest to

	ordering and positioning materials spatially (e.g., from smallest to biggest, from highest to lowest)	smallest… Because maybe some people see it differently, that it's more important to produce more or less." —Teo
Proportional unitizing	Pupils formulating rules to allocate units of measurement to specific quantities of materials to set visual parameters	"We just made a certain, uh, amount for like one hundred, a certain amount for ten, for one, and a certain amount for five." —Mya & Ari
Constructing Spatial Understanding through Embodied Making		
Generating ideas and concepts togive visual-spatial forms to data	Pupils developing preliminary ideas and concepts about the forms or methods they can follow to physicalize data visually and spatially, either in their mind or at the onset of interaction with the materials	"Look, so it's going to be like this, one whole pie, okay? Pie chart…If I wanted to know how much India uses, I could measure how, like, how much uses. I could just, like, into a fraction." —Shay & Otto
Evaluating and adjusting the methods and materials used to physicalize data	Pupils evaluating and adjusting size differences, unit values, spatial ordering, or ways to arrange materials in 2D or 3D as they interact with the materials	"This one has to be filled slightly more…and the plastic it has to be filled exactly the same." —Lia & Ada
Identifying a (mis)match between ideas and crafting skills	Pupils bringing ideas in mind to life through crafting with materials, or realizing the need to modify or give up on some ideas due to limited crafting skills	"So I first thought to make a pyramid, a diorama thing. Something like a, like a 3D thing, yeah. But then I thought that, like, it would take a lot of time, and I'm really not that good at making crafts."—Ravi
Considering viewers' spatial perspectives	Pupils envisioning what their physicalizations will look like when viewed by others from different angles or perspectives	"We can stuff some paper underneath and then put this on top to make it looks like more full. Because in a box you can only see what's on top."—Lia & Ada

Table 2.1 continued

2.4.1 Spatial ways of representation

Quantities-to-sizes mapping

Here is an example of pupils representing value differences in the numerical data they selected by varying the sizes (e.g., length, width, height, volume) of tangible materials either through measuring or estimation (see Figure 2.2).

Figure 2.2 Pom poms of varying sizes—Beau's team

Beau: So, I made with my partner, like, this design to show how much waste our school produces and which type, so it's the bigger they are, the more trash. So like, this is really small, the 10% of the trash is of glass. 14% of the paper, cardboard slash paper. Plastic is also 14%. Compostable waste is a whole 43%. And I guess Compostable Waste was made out of quite a lot of things. Also the pom poms sizes were not enough, so we decided to put multiple pom poms in the balloon to represent the compostable. So, you have to read to get the exact percentage, but you know which type of trash is more than this type of trash if you just look. (After-class interview)

Beau and his teammate decided to use spherical pom poms of varied sizes to represent the differences in the weight (in percentages) of six types of waste produced at their school. By comparing the magnitude of numbers in percentages, they selected pom pom sizes that roughly matched the numerical values through visual estimation. It can be speculated that they either gauged the differences in these spheres' volumes or used the diameters of the spheres as indicators of size. Beau's explanation revealed that he was constantly relating the materials to the numerical values, displaying a clear intention to map quantities of waste to the sizes of pom poms. Interestingly, when they realized that they did not have a pom pom big enough to represent 43%—more than double the amount represented by other pom poms—they decided to combine multiple pom poms not only to represent a much larger amount but also to convey the idea that the compostable waste category consisted of a mix of wastes. In addition, Beau noted that their physicalization made it easy for viewers to understand the data, allowing them to directly perceive which type of waste was produced in larger or smaller quantities.

Spatial ordinal arrangement

Here are two examples of pupils highlighting ordinal relationships in the numerical data they selected by ordering and positioning materials spatially (e.g., from smallest to biggest, from highest to lowest) (see Figure 2.3).

Figure 2.3 Ordered hot air balloon—Pim

Pim: Um, so I'm making a bunch of hot air balloons, and since it shows in kg, the lower the air balloon is, the more kg they carry. Each air balloon is a country. So for the lightest we have, uh, Vietnam up here, which is with 110 kilograms, the lightest of our five balloons. Then comes India with 140 kilograms, also really light, slightly heavier though. And then, we are going down a big drop, Austria has 588 kilograms. So, pretty heavy down here. Then we have Germany and USA.
Researcher: Did you put stuff in the air balloon that made them heavy?
Pim: Um, no. So it's not actually heavier, it's just gonna be placed lower.
Researcher: Why did you make it this way?
Pim: We made it this way because, um, it was really easy to show the way, like, how the weight was, like, weighing them down. Like a hot air balloon, like, it was being weighed down by its weight. (After-class interview).

While the numerical values in the data tables were not ordered, Pim spatially and ordinally arranged the five "hot air balloons" to represent the waste produced per capita in five different countries of his choice. Viewing from left to right, the spatial positions of the balloons indicated that they were ranked from lightest to heaviest. Pim clarified that he conveyed the idea of weight not by making the balloons heavy, but by varying the lengths of the yarn attached to the balloons, thus achieving different positions of the balloons. He assigned visual and spatial properties to the numerical data and re-ordered the data to highlight the ordinal relationship. This was based on the spatial distance between the balloons, with the lower-positioned balloons metaphorically representing heavier weights and the higher-positioned ones representing lighter weights, allowing viewers to understand the spatial relations intuitively. This physicalization exemplified an effective use of three-dimensional space to facilitate viewers' understanding of the data (see Figure 2.3).

Figure 2.4 Ordered clay and sticks—Teo

Teo: So I made this scale, which is representing the five biggest, uh, kilogram countries. USA is the biggest one with 812, Germany, 809 kilograms, and so on down. Um, and each stick, the USA stick is a whole stick, and with each country, which is less kg, it gets smaller each time, as you can see.
Researcher: If you could do it differently, would you change anything?
Teo: Maybe from smallest to biggest, instead of biggest to smallest.
Researcher: Why?
Teo: Because maybe some people see it differently, that it's more important to produce more or less.
(After-class interview)

This is another example of a spatial representation of number magnitude (Figure 2.4). Teo also chose to represent the non-ordered data on the weight of waste per capita in five different countries of his choice. He decided that a whole wooden stick represented the biggest number, and the sticks became shorter as the quantities decreased. Teo's reasoning about number magnitude was aligned with his perception of the sticks' lengths. Interestingly, when asked about alternative approaches, Teo suggested that changing the ordinal arrangement (in descending or ascending order) could highlight different messages. This illustrated his awareness of conveying messages from data to viewers by emphasizing ordinal relationships and manipulating the different possible spatial arrangements of the materials.

Proportional unitizing

Here are two examples of pupils formulating rules for allocating units of measurement to specific quantities of materials to set visual parameters (see Figure 2.5).

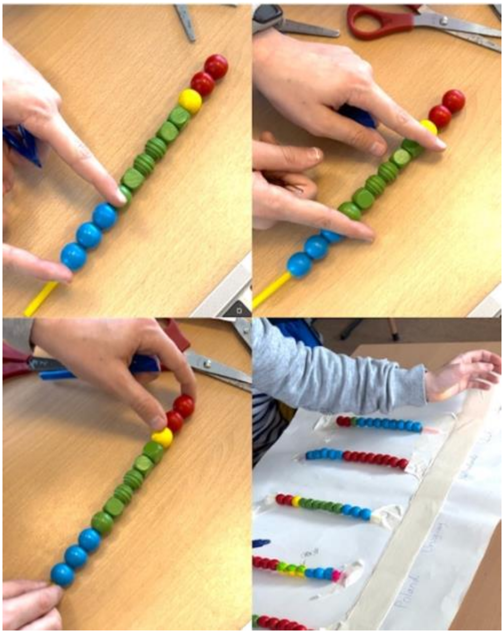


Figure 2.5 Unitized beads—Kane's team

Kane: So, each bead means a different amount of trash that is, that Uraguay has produced. Blue, usually would be 100. So 300. Greens are tens. So they produce three hundred and sixty. Yellows are fives. So, three hundred and sixty-five. And reds are ones. So, it's three hundred and sixty-seven. (In-class observation)

Here, Kane was gesturing to explain his team's decision to use beads of different colors to represent different quantities and assign unit values to colored beads. From his explanation, it became clear that they established meaningful mappings between numerical values in the data and the physical quantities and properties of the beads. To effectively represent large three-digit numbers in 3D with simple and limited materials, Kane and his teammate developed the rule where each blue bead represented 100 kg, each green bead represented 10 kg, each yellow bead represented 5 kg, and each red bead represented 1 kg. This approach demonstrated their proportional reasoning—an important cognitive strategy often used in mathematics that requires spatial reasoning. They then applied this numerical-spatial mapping rule to make stringed beads representing all five countries of their choice (see Figure 2.5).

Ari: Yeah, I think this can be one hundred, this much, this much is one hundred, the way I folded it, and then another one hundred, and then that's the ten. Okay? Good? And then just fold it.
Mya: Oh yeah, and then you can stick it. Now we need to use…
Researcher: Are you measuring it? How are you coming up with that length?
Mya: Um, we just made a certain, uh, amount for like one hundred, a certain amount for ten, for one, and a certain amount for five.
Ari: So maybe like one hundred could be this much. That's smart is it? So, how many centimeters?
Mya: It is… 9 cm. One hundred is 9cm. Ten could be maybe one.
Ari: Okay, yeah, one centimeter because this is exactly, ten is one centimeter. (In-class observation)

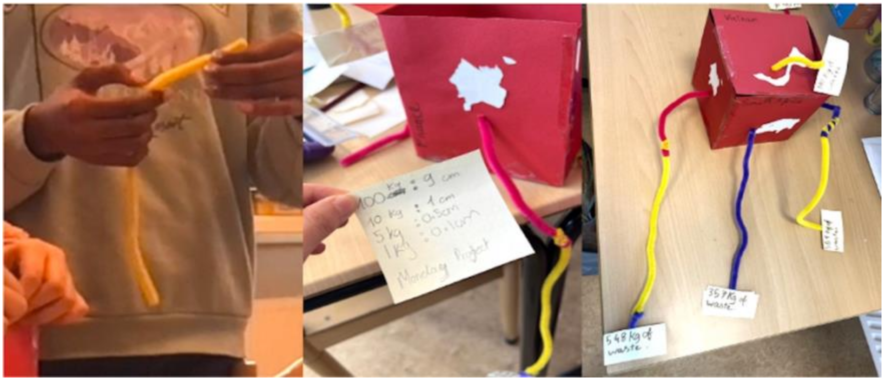


Figure 2.6 Unitized pipe cleaners around a cube—Mya & Ari

Similarly, Ari and Mya discussed how they could represent the large numbers of waste using pipe cleaners (Figure 2.6). Having pipe cleaners at hand, they decided to measure a certain length to represent 100 kg, which turned out to be 9 cm. They quickly deduced that if a 100 kg is represented by 9 cm, 10 kg could be represented by 1 cm. In the figure shown in the middle of the panel, Ari and Mya wrote down the rules that 9 centimeters represented 100 kg, 1 cm represented 10 kg, 0.5 cm represented 5 kg, and 0.1 cm represented 1 kg. It appeared that they were reasoning somewhat proportionally about numerical values from data and the lengths of pipe cleaners. However, they did not realize that, by proportion, 100 kg should correspond to 10 cm instead of 9. Thus, while it was intriguing to see how these children devised rules of unitizing that corresponded to intangible numerical values with visual–spatial properties, it was also evident that their proportional reasoning was imprecise.

2.4.2 Constructing spatial understanding through embodied making

We now proceed to discuss the four themes that place a stronger emphasis on pupils' making process. As the pupils interacted with materials and their surroundings, they developed new understandings of the visual–spatial properties of the materials at hand. They considered the spatial relations between these materials that could depict relations among data variables and also came up with alternative methods to construct their physicalizations.

Generating ideas and concepts to give visual–spatial forms to data

First, we present an example of pupils developing preliminary ideas and concepts about the forms or methods they can follow to physicalize data visually and spatially at the onset of interaction with the materials (see Figure 2.7).

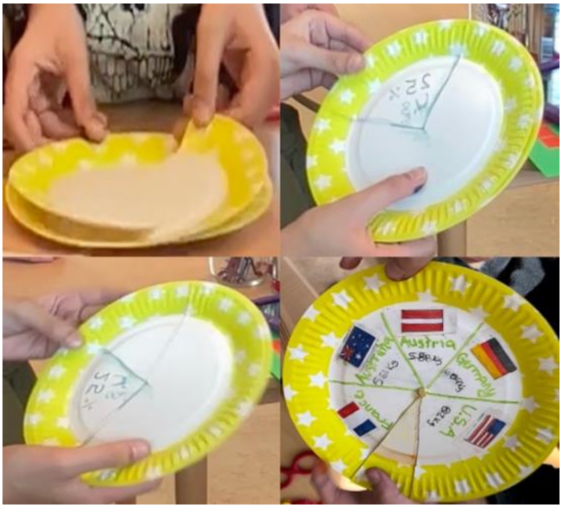


Figure 2.7 Movable pie chart—Shay & Otto

Shay: Look what I made. You pull it. (showing the movable plates) So we could like, um, make a pie chart. We can put the information right here.
Otto: Um, how will we fit five countries on that? I'm confused, um, which paper is which. I do not even know which plate is which anymore.
Shay: Look, so it's going to be like this, one whole pie, okay? Pie chart. And then this is how much… I'm going to have this representing, for example, just give me a pencil, please. Like, this is how much waste India has, okay? If I wanted to know how much India uses, I could measure how, like, how much uses. I could just, like, into a fraction.
Otto: A fraction, yeah, yeah, that would work.
Shay: Yeah. Then we can color it in how much which country to have what. (In-class observation)

In this team, Shay came up with the idea to make movable plates upon seeing the available materials. He joined two plates at the center with a bendable pin and made cuts on both plates, allowing them to spin. With the plates in hand, he thought of converting the data into a pie chart. According to his idea, rotating the plates would reveal the data on waste per capita in different countries sequentially. Although Otto took some time to grasp his teammate's idea, they eventually agreed on converting the data in kilograms to a presentation format suitable for a round plate. They did this by calculating each country's waste per capita as a proportion of the combined per capita values across the five countries. They then engaged in a process of visualizing and estimating. For example, when representing the United States as accounting for approximately 25% of this constructed total, they visualized what one-fourth of the plate's area would look like. This process of estimating the area likely involved spatial thinking steps such as visualizing and comparing areas, and considering the relationship between a portion of the chart and the whole. It is also possible that they considered where the radius of the circle would emerge, visualizing drawing the radius from a certain angle to yield a certain segment of area on the pie chart.

Evaluating and adjusting the methods and materials used to physicalize data

Here are two examples where pupils evaluated and adjusted aspects such as size differences, unit values, spatial ordering, and ways of arrangement while interacting with the materials. Each example showcases different and meaningful approaches to adjusting the materials (see Figure 2.8).

Figure 2.8 Experimenting with cups—Lia & Ada

Lia: I have an idea, so for compost and food we can fill it up until halfway or something (putting the pom poms in the cup to explain the idea) say this is the recycling thing. And fill it up that much. Like, really little, for 5 %. Then, say, for this one (compost), we fill it almost halfway.
Ada: So then, um, how much we fill the plastic?
Lia: We have to fill exactly the same as the paper.
Ada: What if, to make, say, the plastic higher, we can put this on top here, and then we put the lid… And then we can fill that. It's full of plastic. So wait, but is this 14% about?
Lia: This one has to be filled slightly more…and the plastic it has to be filled exactly the same.
Ada: That looks kind of the same.
Lia: Yeah, that's fine, that's fine. Although, this has to be more than this… The tin is 5%…
Ada: So, this is 10% compared to 5%. That's about half.
Lia: Yeah, I think. Look, I had this fun experiment. Let us dump them all into one and see if that makes exactly a full cup. A full hundred percent. I do not think it will because I think it's a bit too much. That makes about a full 100%.
Ada: It's a bit over, right?
Lia: Okay, I'm taking out one of each just to make it slightly smaller. (In-class observation)

As Lia and Ada interacted with the materials, they were actively constructing spatial understandings of the numbers in the data and the data physicalization task itself. Projecting their understandings of the numbers onto the materials, they visually compared and evaluated whether their spatial-numerical mappings were reasonable and whether the spatial representations of number magnitude were accurate. They used the visual feedback from their embodied interactions with the materials to modify the physicalization, likely gaining new insights into how numerical values can correspond to certain heights and volumes.

There are many spatial moments in the episode shown above, such as gauging the space needed to represent 5% and visualizing how much materials they should fill the cup with. They used phrases such as “a bit over” and “slightly smaller” to indicate the process of comparing number magnitudes. Descriptions like “we fill it almost halfway” and “this is 10% compared to 5%. That’s about half” showcased their proportional reasoning of ratios. Overall, it appeared that as they visualized and estimated the sizes of materials they used, they were actively relating number magnitude with their spatial representations—heights and volumes.

Interestingly, Lia came up with an experiment to test the accuracy of their estimations. The idea of checking whether all the materials would add up to 100 percent showed that they were testing their understanding of percentages with real materials. It appeared that the embodied interactions with the materials sparked the idea of experimenting with cups. Precisely because of the use of tangible materials, these pupils could receive direct visual feedback on whether the cup would be full or overfull when they combined everything into one cup. It should be noted that Lia and Ada were focused more on the height of the filled materials rather than the volumes (as the cup they used had a larger surface area on the top than at the bottom). This is expected as they had not yet learned about calculating the volumes of cylindrical shapes like cups (see Figure 2.9).

Figure 2.9 Modifying the balloon stand—Pim

Pim: And after when I was done with all the balloons, I just thought to myself. Oh, well, I guess I could also make like a stand that’s like standing upwards on the table. And then I noticed that it would not just fit all of them and it would not be high enough. So I noticed I could just flip my stand over and it would fit perfectly on the wall. (After-class interview)

Pim’s response revealed that embodied interactions with the materials inspired him to consider an alternative approach for presenting his physicalization. Originally, Pim thought of attaching balloons to a stand on the table, as shown on the top left of the panel. However, he noticed that there was not enough height between the table and the floor. This meant that balloons with longer yarn would simply lie on the floor, failing to convey the differences in their spatial positions. As a solution, Pim decided to flip the stand and affix it high up on the classroom wall, effectively using the classroom space to illustrate the weight differences in waste per capita among countries. Unlike the previous example, where the adjustment was made on the visual mapping process from data to tangible materials, this modification pertained to presentation mapping (Jansen & Dragicevic, 2013) to facilitate viewers' perception of the physicalized data.

Identifying a (mis)match between ideas and crafting skills

The making of data physicalization artifacts required pupils to not only give numbers visual–spatial forms but also to use crafting skills to bring their ideas in mind to life. Occasionally, pupils realized that they needed to modify or give up on some ideas due to their limited crafting skills (see Figure 2.10).

Figure 2.10 From 3D pyramid to 2D collage—Ravi

Researcher: What was the biggest challenge for you?
Ravi: Um, so I first thought to make a pyramid, a diorama thing. Something like a, like a 3D thing, yeah. But then I thought that, like, it would take a lot of time, and I'm really not that good at making crafts. And I would love to get all the supplies and I really do not know how to do it. And after that, representing. Put them in a, uh, pyramid, it's like, tough. And then I think of like a simple way, like other than the pyramid, I thought of a simpler, like, a figure. So I just thought to make a chart out of like, I cut it out some papers and then I stick them on the white piece of paper and I wrote the country's name, which is easy to understand for people and this would take a bit less time. So for an example, I think, uh, Netherlands were the biggest. So like I cut out a big chunk of paper for Netherlands and a small chunk for India because the number was smaller, was like a hundred something. (After-class interview)

Due to a lack of certain crafting skills, Ravi gave up on the idea of making a 3D pyramid shape physicalization and opted for a simple 2D collage representation. In his current physicalization, Ravi varied the side lengths of different strips in the collage, showing that Ravi was giving visual–spatial forms to numerical values and exploring the different properties of two- and three-dimensional space in his mind. Had he pursued the pyramid idea, it would have demanded more spatial thinking. For instance, he would need to decide which of the many side lengths to vary, which taps into his knowledge of this 3D geometric shape. Moreover, making a pyramid from scratch would be challenging, as he might need to sketch a net plan for the pyramid on paper and then assemble it in 3D. While these were merely our conjectures, it is reasonable to hypothesize that crafting a 3D pyramid would have been more spatially challenging for him had he possessed the crafting skills and sufficient time.

Some of these pupils were able to craft spatially complex representations. For example, Mya and Ari (Figure 2.6) built a cube from a net plan and managed to secure the pipe cleaners of varying lengths onto each face of the cube. This process demanded spatial visualization, mental rotation, and possibly prior knowledge about crafting a paper cube from scratch and making knots on paper. For another example, Shay (Figure 2.7) used simple materials to create a movable design that allowed for spatial transformation. Nevertheless, data physicalization tasks can still be challenging for some pupils, given the limited crafting skills, tight class schedule, and restrictions in available materials.

Considering viewers' spatial perspective

Here is an example of pupils envisioning what their physicalizations would look like when viewed by others from different angles or perspectives. This conversation is part of an earlier discussion between Lia and Ada (Figure 2.11).

Figure 2.11 Cups from different perspectives—Lia & Ada

Lia: Say we do not have enough of glass. We can stuff some paper underneath and then put this on top to make it looks like more full. Because in a box you can only see what's on top.
Ada: Oh and we can put a mark on the front to say how much of it is inside, but you could also look inside to see like how full it is. (In-class observation)

Lia and Ada had a limited number of plastic lids available, which were insufficient to represent the quantity of plastic waste. Interestingly, Lia suggested placing other materials underneath and simply covering the top with the two plastic lids they had. She reasoned that viewers could only see what was on the surface when looking from the top. Ada quickly built on this idea, suggesting that they could label the percentages on the cup's outer wall to facilitate viewers' understanding from the side. Their discussion revealed that they were envisioning viewers actively perceiving their physicalization from different angles. They developed this solution not just to tackle the problem of limited materials but also to ensure that viewers receive accurate data information, whether viewing from above or from the side.

2.4.3 Self-report from pupils

22 of the 37 pupils completed the self-report survey individually. The remaining pupils did not complete the survey, possibly because they missed the instruction, had not finished their artifacts, or misunderstood the instruction, so that only one child from a team filled out the survey. Pupils gave Likert-scale ratings from 1 to 5—1 indicating "I'm not doing it yet/Does not sound like me yet", 3 indicating "I did it sometimes/Sounds somewhat like me", and 5 indicating "I did it often/Sounds very much like me." Among the respondents, 95.5% of the pupils gave a rating of 3 or above for the statement "Before making, I thought about how the visualization would look"; 81% of the pupils gave a rating of 3 or above for the statement "During making, I checked to see if the visualization represents the data well"; 95.5% of the pupils gave a rating of 3 or above for the statement "I can explain the meaning of the data to others using the visualization"; 100% of the pupils gave a rating of 3 or above for the statement "I can understand the data better now"; and 72.7% of the pupils gave a rating of 3 or above for the statement "I have more ideas for what I want to design for the design challenge."

In summary, all pupils who filled out the survey reported an enhanced understanding of the data about waste. A majority were capable of visualizing what their representation would look like before making and articulating the meaning of their data through their data physicalizations. A number of pupils reported that they occasionally or frequently checked to see if the representation matched the data well during the making process. Additionally, some of these pupils reported that the data physicalization activity inspired their design ideas for the Circular Economy design theme.

2.5 Discussion and conclusion

In this case study, pupils utilized their spatial thinking alongside their numerical understanding and crafting skills to devise a variety of methods for mapping numbers to visual–spatial forms. Specifically, they made creative and effective use of two- and three-dimensionality, resulting in data physicalization artifacts that incorporated a range of materials, sizes, and configurations, depicted ordinal or proportional relationships, and engaged viewers to perceive these artifacts from different angles. At its core, pupils' construction of data physicalizations revealed insights into their process of encoding different dimensions of information from the data and transforming non-spatial data into tangible objects with visual–spatial properties and relations.

2.5.1 Using visual and spatial properties to encode data—spatial ways of representation

To address the first research question, our analysis of pupils' data physicalization artifacts, conversation, and making processes triangulated evidence that the pupils used spatial thinking as they mapped numerical data onto tangible materials. Among the different approaches to physicalization, we identified three ways of visual mapping that potentially recruited spatial thinking and were repeatedly observed among the pupils in our study.

Quantities-to-size mapping, the technique of representing value differences in numerical data by varying the sizes of tangible materials, is commonly seen in data physicalizations made by adults (e.g., Vande Moere & Patel, 2010; Jansen & Hornbæk, 2015; Hurtienne & Reinhardt, 2017; Perin, 2021). This technique was also reflected in all pupils' artifacts mentioned above (Figures 2.2–2.11). The cognitive strategy of mapping intangible numbers to tangible entities is important in mathematics and requires pupils' understanding of quantities and proportional reasoning (Lamon, 1996). This practice recruits spatial thinking, evident in our results where pupils corresponded numerical values with indicators of size, such as a certain length, width, area, or volume of tangible materials, either through visual estimation or precise measurement. Moreover, to illustrate a comparison of number magnitude in data, pupils leveraged their knowledge of ratios and proportions, for instance, by spatially scaling the material's size to represent data values that doubled or tripled (Figures 2.2 & 2.8).

Proportional unitizing is a reasoning strategy critical to math learning and frequently employed in data visualization and physicalization by novices and experts to map numerical data to visual elements (Huron, Carpendale et al., 2014). In a study by Huron, Jansen, and Carpendale (2014), adult participants with little experience in information visualization were given unit tokens to conveniently construct their physical visualizations. From our results, we noticed that pupils spontaneously thought of using unit rates (Figures 2.5 & 2.6). Multiple teams developed and adhered to proportional unitizing rules, demonstrating thoughtful consideration of the materials' visual–spatial properties and adeptness at representing large three-digit numbers with materials of limited sizes and shapes. For other pupils, this process may have been intuitive rather than rule-based, indicating an implicit understanding of proportional unitizing.

An additional way of representation that required spatial thinking and was evident in pupils' artifacts was using spatial ordinal arrangement to highlight meaningful ordinal relationships within the data. This technique is frequently seen in data visualization and physicalization practices of adults (e.g., Ware, 2004; Vande Moere & Patel, 2010). Similarly, in our results, several teams or individual pupils ranked the materials they used to represent data values in either ascending or descending order (Figures 2.3 & 2.4). These arrangements are meaningful to the discussion of spatial thinking because the data tables they received were not sorted in specific orders. As pupils positioned materials in a certain order in the 3D space, they were likely contemplating the relationships between data variables and reasoning how these relationships could be conveyed through spatial proximity and distance.

2.5.2 The interplay between spatial thinking and embodied making

Our second research question was: How does the embodied process of making data physicalization stimulate pupils' spatial thinking? We observed that thinking along and interacting with tangible materials meaningfully stimulated pupils' spatial thinking and offered them opportunities to transform numbers into tangible representations with visual–spatial properties.

The theme, generating ideas and concepts to give visual–spatial forms to data, illustrates how pupils developed ideas for diverse ways of physicalization through interacting with the materials (e.g., Figures 2.3, 2.6–2.9). For example, manipulating pipe cleaners prompted pupils to correspond certain lengths with certain weights (Figure 2.6). Similarly, having movable plates at hand gave pupils the idea of converting data in kilograms into percentages and mapping them onto respective areas on a pie chart-like plate (Figure 2.7). Our observations revealed that all teams or individual working pupils were able to give visual–spatial forms to data, either in 2D or 3D, through active interactions with the materials. This resonates with prior studies suggesting that creating visual–spatial representations facilitates problem exploration as well as perception organization (Tytler et al., 2013). It may also be explained by earlier findings, which indicated that tangible materials and manipulatives function as a “mental tool view” (Chao et al., 2000) and that interactions with physical materials elicit mental imageries that can “guide and constrain their thinking for problem solving” (p. 285).

The theme, evaluating and adjusting the methods and materials used to physicalize data, is exemplified by pupils experimenting with both visual mapping and presentation mapping. Pupils’ frequent hand movements around the materials helped them test the accuracy of their spatial understanding of number magnitude (Figure 2.8) and uncover the optimal spatial layout of their physicalization (Figure 2.9). It appeared that making tangible representations of data allowed them to visually compare and evaluate the appropriateness of their representations, and any adjustments made may have updated their mental representations (Huron, Jansen, & Carpendale, 2014). These complex embodied thinking processes echo Schön’s (1983) notions of “knowing in action” and “reflecting in action.” They also align with Kelly et al.’s (1987) depiction of the ongoing interaction between mind and action in the design processes of 15-year-olds, where “making activities interact with cognitive activities as the manifestations of our ideas allow us to think more deeply about their implications” (p. 13). In addition, our results under this theme appear to be in line with the findings that spatial motor movements serve as a manifestation of the underlying spatial perception and cognitive reasoning (Maresch & Sorby, 2021) and elicit visual feedback to support mental simulation (Frick et al., 2009).

The theme, considering viewers’ spatial perspectives, aligns with the suggestion that designers need to envision viewers taking an active rather than passive stance when interacting with physicalized data (Jansen et al., 2015). As depicted in Figure 2.11, pupils were contemplating how their physicalization could convey a coherent message about quantity to viewers from either a top or side view. This reasoning process is somewhat analogous to solving spatial visualization problems, where students are often tasked to imagine how a set of blocks looks when viewed from the top or the side, or determine which blocks will be visible or hidden (e.g., the Middle Grades Mathematics Project Spatial Visualization Test by Ben-Chaim et al., 1988). However, unlike solving a paper-and-pencil task on visualizing the different view plans of blocks, the pupils in our study were using their spatial visualization skills to solve an authentic problem that arose from the many possible arrangements of tangible materials. It allowed for (1) open-ended ways to formulate solutions, and (2) hands-on experimentation with direct visual feedback from the materials. For example, they could check to see if placing plastic lids above other materials would give the visual effect of a fuller cup.

To summarize the themes discussed above, the converging evidence from pupils’ data physicalization artifacts, conversation, and making processes indicated that pupils were constructing spatial understandings through the embodied making process of data physicalization, and that their spatial thinking was grounded in their embodied interactions with the materials and the classroom space. Spatial thinking skills are predominantly trained with the goal of having students perform manipulations of mental imagery in mind. However, our findings, consistent with prior research (e.g., Casey et al., 2008; Frick et al., 2009; Link et al., 2013; Burte et al., 2017; Maresch & Sorby, 2021), highlight the important role of embodied experiences in spatial thinking. Our society tends to categorize human activities strictly into physical or intellectual work (Pallasmaa, 2017). Yet we should not always view bodily motions and cognitive capacities as entirely separate agencies. Many examples from our results indicated that had these pupils been limited to creating infographics in their minds or on paper, they might not have actively envisioned numbers with visual–spatial properties.

Despite the observed benefits of embodied making experiences on pupils' construction of data physicalization and their use of spatial thinking throughout the activities, we are mindful of the ongoing debate regarding the role of embodied experience in supporting pupils' spatial thinking development. On one hand, there is the perspective that bodily experiences and mental activities complement each other (Wilson, 2002; Pallasmaa, 2017) and that embodied experiences are crucial for spatial thinking development (Frick et al., 2009; Link et al., 2013). On the other hand, there is concern that the use of hands-on manipulatives may diffuse the emphasis on honing pupils' abilities to visualize, plan, and solve problems in their minds without external aids (Newcombe, 2017). More research is needed to fully understand how embodiment influences pupils' spatial thinking and to determine under which conditions it nurtures or impedes the development of certain spatial skills.

Lastly, we observed a display of spatial creativity, defined as "the ability to create new and rich 2D and 3D ideas for volumetric forms that consist of spatial configuration, organization, and the spatial relationships of the components" (Suh & Cho, 2020, p. 3). For example, the pupils in Figure 2.6 attached traced shapes of countries to the faces of a cube, in addition to presenting unitized data information on waste per capita in these countries. In Figure 2.7, the pupils created movable plates that allow spatial transformation and interaction with viewers. Unlike visualizing data in a 2D or virtual 3D environment, a physical environment promotes tactile exploration and interactions, allowing the maker to embed the meaning of data in creative ways and the audience to experience data in unexpected ways (Vande Moere & Patel, 2010). We should note that not all of the representations were novel forms, as some resembled conventional bar or line graphs (e.g., Figure 2.4). Nevertheless, the diversity of artifacts produced by these pupils still indicated their application of everyday creativity to develop what seemed original and valuable to them (Craft, 2001; Cremin et al., 2018).

Overall, pupils of this age seem to be able to meaningfully engage in data physicalization activities. Pupils exercised their spatial and numerical understandings during this brief, versatile classroom activity, which has the potential to be implemented in design and technology education as well as in broader STEM contexts. Based on these findings, we will move forward to discuss the practical implications of this study for educators and researchers.

2.6 Implications and future research

The data physicalization activity discussed in this case study was integrated into a design module and took place in a formal learning environment. This activity effectively directed pupils towards thinking about waste-related issues, devising methods to communicate data information to others, and subsequently generating ideas on designing through reusing and repurposing waste. Our approach highlights the potential of using data physicalization as a problem-exploration or storytelling tool in design education.

In line with previous findings that data physicalization can be a valuable tool for learning in the classroom (Willett & Huron, 2016), our findings have several implications for STEM education. Firstly, this qualitative examination paves the way for the development of embodied, educational interventions targeting pupils' spatial-numerical understandings. Secondly, given the low-tech and versatile nature of data physicalization activities, coupled with the importance of data literacy across disciplines, there is potential to weave data physicalization activities into STEM, arts, or social science classrooms to facilitate subject knowledge comprehension and enrich the learning experiences with embodied elements. Furthermore, resonating with observations from Bae et al. (2022), we noticed that data physicalization provided pupils with ample opportunities for hands-on experimentations, such as seeing the change in height and volume as they fill in materials and conveniently remove materials to achieve the desired quantity (Figure 2.8). Hence, data physicalization may be a promising tool for experiential learning in STEM.

There is room for improvement regarding how the data physicalization activity was organized. Firstly, additional scaffolding may be needed to help pupils better understand data and the physicalization task. Generating visual–spatial representations based on verbal information requires thoughtful selection and integration of information (Van Meter et al., 2006). The use of tangible materials poses further challenges to pupils' ability to integrate information, taking into account the visual, spatial, and physical properties of the materials used. Therefore, it could be beneficial to guide pupils in identifying key information in the dataset, exploring the properties of the materials, and reflecting on effective or alternative ways of representing the data. For instance, the movable pie chart group (Figure 2.7) could benefit from teacher prompts encouraging consideration of different representational formats better suited to comparing values across countries.

Secondly, due to scheduling challenges, only pupils in the afternoon class had the opportunity to present their physicalizations to the class. Through practice, we noticed the importance of presenting and receiving feedback, as these enabled pupils not only to create but also to critically evaluate their own work and that of their peers. For example, feedback nudged pupils to think about how comprehensible their physicalization is to others. By doing so, educators can better tailor these activities to nurture pupils' data literacy, visual literacy, and representational competency. Therefore, in future data physicalization practices, it would be beneficial to provide opportunities for self- and group-reflection, as well as peer feedback. Finally, the pupils involved in our study come from middle to upper socioeconomic backgrounds, with some having parents in STEM professions. This background might have given them greater exposure to STEM-related knowledge and skills, potentially aiding their data physicalization construction. Future research should consider organizing data physicalization activities with pupils from varied backgrounds and different levels of data and visual literacy.

It is important to note that pupils' ability to create elaborate and spatially complex physicalization artifacts is constrained by their crafting skills. This limitation is exemplified in the theme, identifying a (mis)match between ideas and crafting skills. Oftentimes, the three-dimensional, spatial ideas that were vivid in their minds were compromised due to insufficient crafting skills. Future research is needed to investigate whether training in crafting and making skills can benefit pupils' spatial making processes.

Lastly, future research should investigate how pupils with varying spatial ability levels may benefit differently from data physicalization activities. In the current study, our analysis did not explore the dynamics between pupils working in pairs. Yet, it is conceivable that a child with higher spatial ability might adopt more spatial methods for data physicalization or be more adept at visually estimating, measuring, or crafting. Prior research indicated that external representation and physical manipulation often reduce the cognitive load required for mental computations (Kirsh, 2009). However, some suggested that specific external representations (e.g., on-screen 3D models) may only benefit students with higher spatial abilities while inducing more cognitive load on students with lower spatial abilities (Huk, 2006). Future studies should explore how the use and fabrication of tangible visualizations may influence pupils with varying spatial ability levels in distinct ways. While our study provided qualitative insights into how tangible visualization activities engage spatial thinking, it is necessary for future quantitative or mixed-methods studies to investigate how these activities can be structured to enhance pupils' spatial ability.

CHAPTER 3

Design-by-analogy: Pupils' spatial thinking in an open-ended, biomimicry design challenge

CHAPTER 3
Design-by-analogy: Pupils' spatial thinking in an open-ended, biomimicry design challenge[3]

Spatial thinking is ubiquitous in design. Design education across all age groups encompasses a range of spatially challenging activities, such as forming and modifying mental representations of ideas, and visualizing the scenarios of design prototypes being used. While extensive research has examined the cognitive processes of spatial thinking and their relationships to science, technology, engineering, and mathematics learning, there remains a knowledge gap regarding the specific spatial thinking processes needed for open-ended problems, which may differ from those assessed in closed-ended, analytical spatial tasks. To address this gap, we used educational design-based research to develop a nature-inspired, design-by-analogy project and investigate the spatial thinking processes of young, novice designers. 16 pupils from an international school in the Netherlands participated in this five-week design project. Multimodal evidence from classroom recordings and pupils' design works were triangulated to offer insight into the key spatial thinking processes involved in their creation of nature-inspired, analogy-based design prototypes. Results revealed spatial thinking processes that might not align with those assessed in conventional spatial tests and may be unique to design or open-ended problem-solving. These processes include abstracting spatial features to infer form-function relationships, retrieving a range of relevant visual information from memory, developing multiple possible analogical matches based on spatial features and relationships, elaborating and iterating on the design concepts and representations to make creative and suitable solutions for the design challenge, as well as visualizing design prototypes in practical usage scenarios. By highlighting the nuanced differences between spatial thinking in open-ended, divergent thinking tasks and conventional spatial tasks that demand single correct solutions, our research contributes to a deeper understanding of how pupils utilize spatial thinking in design and open-ended problem-solving contexts. Furthermore, this case study offers practical implications for scaffolding pupils' analogical reasoning and nurturing their spatial thinking in design education.

[3] *This chapter includes text that has previously appeared in the following source:*

Zhu, C., Klapwijk, R., Silva-Ordaz, M., Spandaw, J., de Vries, M. J. (2024). Investigating the role of spatial thinking in children's design ideation through an open-ended design-by-analogy challenge. International Journal of Technology and Design Education.

3.1 Introduction

The ability to think spatially and deal with spatial information requires "searching the visual field, apprehending the forms, shapes, and positions of objects as visually perceived, forming mental representations of those forms, shapes, and positions, and manipulating such representations" in one's mind (Carroll, 1993, p. 304). These cognitive processes are often externalized through spatial representations like diagrams, sketches, and models (Gagnier et al., 2017; Novick et al., 1999; Stieff et al., 2005) as well as depictive or dynamic hand motions (Clement, 2008; Ehrlich & Levine, 2006; Pallasmaa, 2017). Together, these cognitive processes play an important role in the learning and practicing of science, technology, engineering, arts, and mathematics (STEAM) disciplines, including design (e.g., Buckley et al., 2018, 2022; Hegarty, 2014; Kell et al., 2013; Lin, 2016; Wai et al., 2009).

Over the past decades, a number of studies have confirmed that spatial thinking abilities can be trained and improved (e.g., Cheng & Mix, 2014; Hawes et al., 2017; Lowrie et al., 2017; Sorby, 2009; Uttal, Meadow, et al., 2013). Since spatial thinking is not used in a vacuum but often along with content knowledge (Atit et al., 2020; Hegarty et al., 2007; Ormand et al., 2014), it has been proposed that psychological and educational interventions should be designed to enhance not only individuals' spatial abilities but also their problem-solving performance in real-world learning tasks (e.g. Atit et al., 2020; National Research Council, 2006; Newcombe, 2017; Uttal & Cohen, 2012; Uttal, Miller & Newcombe, 2013; Zhu, Leung et al., 2023). To understand and support students' use of spatial thinking in authentic learning contexts, research has been dedicated to integrating spatial practices in day-to-day classrooms, such as in mathematics (Hawes et al., 2017), chemistry (Stieff et al., 2012), geology (Ormand et al., 2017), and engineering (Julià & Antolì, 2018; Sorby & Baartmans, 2000). However, there has been relatively limited exploration of developing spatial thinking through Design and Technology (D&T) education.

Engineers, scientists, and designers regularly engage in design practices that involve integrating interdisciplinary knowledge and skills to devise purposeful solutions to authentic problems (Klapwijk & Stables, 2023). Spatial thinking and the use of spatial representations are prevalent in design practices, especially when the designed products and processes have a material nature. As characterized by Schon and Wiggins (1992), "A designer sees, moves and sees again" (p. 135). Designers' exploration of the design problem space and solution space is visualized through their minds' eyes and often externalized through multimodal representations, including verbal elaborations, sketches, gestures, and the making of tangible or virtual prototypes (Kavakli & Gero, 2001, 2002; Schon & Wiggins, 1992). Attending to the visuo-spatial features in their design artifacts, such as the shapes, sizes, spatial relations, or spatial arrangements, may in turn lead to 'unexpected discoveries' about the design problems and possible creative solutions (Suwa et al., 2000). Unlike tasks commonly studied by cognitive scientists, design problems do not have a single correct answer (Cross, 2006), and the process of developing design solutions has, in essence, a divergent nature (Guilford, 1967). Moreover, designers often need to envision and create things that do not yet exist by visualizing the new products both in their minds and through materials. As a result, it is reasonable to speculate that the cognitive spatial processes used in design ideation may differ, to some extent, from those employed and assessed in widely used spatial tasks with predetermined answers (e.g., Guay, 1977; Shepard & Metzler, 1971; Vandenberg & Kuse, 1978).

To generate creative solutions to open-ended design problems, designers employ various techniques to stimulate divergent thinking, such as re-representing or reframing problems (Zahner et al., 2010), examining pictorial examples (Purcell & Gero, 1996), and using design-by-analogy (Fu et al., 2015; Hey et al., 2008; Linsey et al., 2012; Qian & Gero, 1996). In the current study, we aim to understand how pupils utilize spatial thinking in an open-ended, design problem-solving situation that requires divergent thinking and analogical thinking. To investigate this, we developed a biomimicry design project, where pupils draw inspiration from various forms in nature, map form-function relationships observed in nature onto their own designs, and create spatially complex design-by-analogy prototypes. We expect that gaining insights into pupils' spatial thinking in this particular context will help identify spatial thinking processes that are unique to open-ended problem-solving.

3.2 Literature review

3.2.1 Developing spatial thinking in authentic learning contexts

Spatial ability has been a subject of extensive research in cognitive science, developmental psychology, and educational science (e.g., Gilligan et al., 2020; Hawes et al., 2017; Lowrie et al., 2019; Mix et al., 2021; Newcombe, 2017; Uttal, Miller & Newcombe, 2013). The association between individuals' spatial abilities and their academic performance has been investigated in disciplines such as science (Stieff et al., 2012), technology (Buckley et al., 2022), engineering (Hsi et al., 1997; Sorby, 2009), and mathematics (Hawes et al., 2015). There is a general consensus that higher spatial abilities are linked to better performance in these subjects as well as an increased likelihood of pursuing careers in these domains (Shea et al., 2001; Wai et al., 2009). Since spatial thinking skills develop from a young age and throughout childhood (e.g., Newcombe & Huttenlocher, 2000), providing school-age pupils with opportunities to engage in spatial activities within formal, informal, and non-formal learning environments has the potential to enhance their spatial abilities (Newcombe & Frick, 2010; Newcombe & Stieff, 2012).

Despite the considerable amount of lab-based research conducted to develop K-12 students' spatial thinking skills (e.g., Cheng & Mix, 2014; Gilligan et al., 2019; Xu & LeFevre, 2016), increasing effort has been made to situate spatial training in authentic learning contexts (e.g., Akayuure et al., 2016; Hawes et al., 2017; Lowrie et al., 2019; Taylor & Hutton, 2013; Zhu, Leung et al., 2023), which means situating learning in meaningful real-world situations (Herrington et al., 2014). By integrating spatial training or spatial activities into authentic learning contexts, findings from cognitive science research can be translated into educational practices. For example, Lowrie and colleagues (2019) developed a spatial thinking program for upper primary school pupils, where teachers provided instruction aimed at supporting pupils' spatial understanding of geometry knowledge. Following the three-week program, pupils showed improvements in spatial visualization, spatial orientation, and mathematics performance.

Most existing school-based spatial interventions, however, have primarily focused on mathematics education, which emphasizes using analytical thinking and convergent thinking to arrive at a single correct solution to a problem. In contrast, less attention has been given to design education, which places a strong emphasis on divergent thinking. The potential of engaging pupils in spatial thinking through design education warrants in-depth investigation. Yet before spatializing design education, it is essential to understand how pupils use spatial thinking when generating and implementing their design ideas, as the cognitive processes involved may differ from those used in solving math problems with predetermined solutions.

3.2.2 Spatial thinking in open-ended problem solving—with a focus on design

Our ability to bring our imaginings to life is a capacity that has shaped the world for centuries (Klapwijk & Stables, 2023). Since the eighties, the teaching of Design and Technology practices in primary and secondary education has gained popularity in countries such as the United Kingdom (Kimbell et al., 1996), the Netherlands (Raat & de Vries, 1986), and more recently under the term integrated STEAM or maker education in the United States, China, Singapore, and Australia (Blikstein, 2018; Cook & Bush, 2018; Zhan et al., 2022). Design problem-solving encompasses various processes such as problem identification and framing, user research, context mapping, identifying constraints and wishes, generating prototypes that represent design ideas in any medium (Houde & Hill, 1997), and testing and iterating on prototypes. Through iterative design processes, both professionals and novices can generate innovative and socially relevant solutions. For instance, professionals designed approaches to harvest solar energy that were inspired by leaves in nature (Benyus, 1997), while primary school pupils were able to design creative and functional ways to represent time without directly displaying it (Klapwijk & Stables, 2023).

While forming a mental representation of a tangible object in our daily life may sound simple, it demands various spatial thinking processes (Lane & Sorby, 2022), such as visualizing and memorizing the details. The development and realization of a design idea involve even more complex spatial thinking processes (Finke et al., 1992; Kavakli & Gero, 2001, 2002; Kosslyn, 1980; Purcell & Gero, 1998; Schon & Wiggins, 1992; Suwa & Tversky, 1997; Suwa et al., 1998; Williams et al., 2010), such as employing mental imagery and making use of spatial relations. A common creative thinking process consists of retrieving visual patterns from memories, mentally re-arranging, reassembling, synthesizing, and transforming what was known into something new, functional, and meaningful (Finke et al., 1992, p. 20–21), and visualizing the scenario when one explores the functions of the new design (p. 25). Furthermore, sketching and constructing three-dimensional prototypes allow designers to share their ‘perception of the space’ with users (Allen, 2010). It is important to note that not only trained designers but also lower secondary school students have demonstrated the use of spatial thinking when tackling open-ended engineering design tasks (Ramey & Uttal, 2017), including discussing spatial information, gesturing to represent static or dynamic spatial arrangements, and using analogical thinking to make sense of spatial properties and relations.

Research in design education, cognitive psychology, and neuroscience has endorsed the important role of spatial thinking in creative problem-solving (Aziz-Zadeh et al., 2013; Benedek et al., 2014; Chang, 2014; Finke et al., 1992; Kell et al., 2013; Muller, 1989; Suh & Cho, 2020). Conversely, a lack of necessary spatial skills may hinder the creation, comprehension, and manipulation of spatially complex designs (Sorby, 2009; Suh & Cho, 2020). Given its interdisciplinary and embodied nature, design as a learning process aligns well with the need to cultivate spatial thinking skills necessary for real-world problems. However, there has been a limited connection between cognitive research on spatial thinking and learning activities in engineering education (Ramey & Uttal, 2017) or in Design & Technology classrooms. As a result, design educators may lack the empirical understanding to identify and support spatial practices within design classrooms.

When faced with well-defined problems that have only one correct solution, such as determining the melting point of ice or selecting the rotated version of an image from multiple choices, one typically relies on convergent thinking. Design problems, however, are ill-defined, meaning that they are open-ended and can have many possible and innovative solutions (Cropley, 2006; Cross, 2006; Finke et al., 1992; Guilford, 1968; Purcell & Gero, 1996). While both divergent and convergent thinking are essential for developing novel and valid ideas (Cropley, 2006; Goldschmidt, 2016; Schut et al., 2020; Zhu et al., 2019), designers use divergent thinking extensively to explore many possible directions before ultimately converging on and evaluating one or several desired solutions (Finke et al., 1992; Goldschmidt, 2016; Guilford, 1968).

Research in neuroscience and cognitive psychology has also highlighted considerable differences between divergent and convergent thinking modes (Gabora, 2010). Divergent thinking requires a broad exploration space to generate a wide range of possible associations, while convergent thinking focuses on identifying and analyzing the cause-and-effect within a certain exploration space (Gabora, 2010). In mathematics education, which has traditionally emphasized using analytical thinking to derive the correct answer, some researchers have argued for the importance of cultivating creative and divergent mathematical thinking through open-ended problems (Becker & Shimada, 1997; Kwon et al., 2006). For instance, Kwon and colleagues (2006) conducted twenty learning sessions featuring open-ended mathematics problems with first-year middle school students. They found that traditional convergent mathematics problems offered students limited opportunities to explore and express different possible solutions. In contrast, open-ended problems allowed students to "try and find their own answers to the problems within their own scope and range of abilities" (p. 57), not only assessing their subject knowledge but also fostering their creativity and divergent thinking.

Existing research on spatial ability has predominantly focused on assessing or improving students' abilities to solve convergent, analytical spatial tasks, resulting in a gap in understanding how spatial thinking is utilized in open-ended tasks, which rely heavily on divergent thinking. Prior research suggests that designers may employ problem-solving approaches distinct from those in other disciplines, such as science (Cross, 2006). Lawson (1979) observed differing approaches between architectural design students and science students when facing a spatial problem. When tasked with devising an optimal spatial arrangement using three-dimensional blocks, science students often began by examining the rules underlying the problem to establish criteria for possible solutions. In contrast, architectural design students experimented with various solutions to identify the best fit and, in doing so, gained an understanding of the underlying rules. Given these findings, it is plausible that the spatial thinking processes involved in design may differ from those observed in other disciplines (e.g., science and mathematics) or those assessed in widely-used spatial tests, such as the mental rotation test (Vandenberg & Kuse, 1978) and the mental paper folding test (Shepard & Feng, 1972).

Potential relationships between spatial thinking and divergent thinking may be inferred from prior investigations into the relationship between fluid intelligence and divergent thinking by Nusbaum and Silvia (2011) and Beaty et al. (2014). Both studies included paper-folding tasks—which require participants to use spatial thinking to mentally visualize paper being transformed and altered—as one of the three measurements of fluid intelligence. Moreover, Nusbaum and Silvia's (2011) study incorporated a cube comparison task, where participants needed to use spatial visualization and spatial orientation to make decisions based on patterns shown on different cube faces. Results from both studies indicated that higher levels of fluid intelligence were associated with higher-quality divergent thinking responses. This correlation is particularly intriguing given that spatial ability was used as one of the proxies in assessing fluid intelligence. The link between spatial thinking and divergent thinking certainly merits further investigation. To offer a fresh perspective beyond existing quantitative examinations, our case study seeks to understand the role of spatial thinking in design ideation—an open-ended problem-solving process that demands divergent thinking. Considering that numerous prior studies were conducted in lab-based or test-based settings, and that authentic problems appear to be better indicators of real-world creativity than standard divergent thinking measures that lack real-world relevance (Okuda et al, 1991), our investigation is situated in an authentic design learning context.

3.2.3 Thinking and designing with visual analogies

Designers make use of a variety of brainstorming techniques and reasoning tools, including analogy (Fu et al., 2015; Goel, 1997; Hey et al., 2008; Linsey et al., 2012; Moreno et al., 2015; Qian & Gero, 1996). Thinking analogically means using knowledge about one situation to inform a novel situation (Gick & Holyoak, 1983; Gentner, 1987; Sternberg, 1977)—in essence, transferring insights from the source to the target. This form of reasoning is central to the cognitive processes behind creative innovations (Chrysikou, 2014; Green et al., 2012) and is frequently employed by designers as a problem-solving and innovation strategy (Ball & Christensen, 2019; Beveridge & Parkins, 1987; Boden, 2001; Casakin & Goldschmidt, 1999; Daugherty & Mentzer, 2008; Gero, 1996; Goel, 1997; Goldschmidt, 1994, 1995). A classic example of design-by-analogy can be seen when engineers in Japan drew inspiration from the beak of the kingfisher bird to redesign the front of a bullet train, thereby enhancing its speed, reducing noise, and improving energy efficiency. Through analogical thinking, designers can transfer information from one domain to another to creatively explore solutions to open-ended problems (Finke et al., 1992). Visual analogies, in particular, have been found to stimulate the generation of novel solutions to design problems (Goldschmidt, 2001; Wilson et al., 2010) and are especially beneficial for novice designers (Casakin, 2004; Casakin & Goldschmidt, 1999).

A number of studies have revealed the important role spatial thinking plays in comprehending and working with visual analogies. “Higher-order visual-spatial thinking is inherently analogic,” as Mathewson (1999) stated, and is “based on comparisons of mental representations” (p. 38). While some analogies can be reasoned through verbal representations, others rely heavily on comparing and making sense of visual mental representations (Beveridge & Parkins, 1987), which means mentally recreating the forms of actual objects or events, any spatial relations pertaining to them, and any dynamic processes happening to or in them (Clement, 2008; Finke, 1989; Kosslyn, 1980). These representations allow one to reason about spatial relations (Beveridge & Parkins, 1987; Huttenlocher, 1968), visualize problem context (Suwa et al., 2000), and generate hypotheses about extreme cases (Clement, 2008).

Through think-aloud interviews, Clement (1981, 1986, 1988, 2009) discovered that when participants generated analogies to solve technical problems, much of their thinking appeared to be spatial, such as visualizing how springs were stretched based on diagrams of coiling springs. For instance, one participant spontaneously came up with the analogy that “a spring is nothing but a rod wound up” and analyzed in his mind how rods might stretch similarly to springs under force (Clement, 1981, p. 3). During these thought processes, Clement (1988, 2008) observed that the participant mentally visualized the problem, compared mental models of a stretching spring to other objects that shared similar key features, imagined possible physical transformations, and determined whether the inferences drawn from one situation could be validly applied to another. Clement further noted that using analogies in problem-solving often entails the spatial reasoning process of "mentally performing imaginative spatial transformations such as deforming, cutting, and reassembling objects in novel ways" (2008, p. 1).

Recent research has shed light on how analogy can serve as a tool for supporting spatial understanding and thinking. Gentner and colleagues (2016) conducted a study at the Chicago Children's Museum with a group of 6-to-8-year-old children. They found that a brief training focusing on analogical spatial structures enabled the children to learn about an important spatial concept that diagonal structures provide stability. In another study, Yuan and colleagues (2017) examined the spatial reasoning and analogical reasoning performance of 3-year-olds. The children were tasked with finding hidden items in 3D rooms based on 2D maps. Through four experiments, the researchers found that children showed a better understanding of the maps and performed more successful spatial searches when there were consistent relational alignments between the source situation and the target situation.

The studies discussed above demonstrated how spatial thinking is utilized to mentally represent, compare, analyze, and draw inferences from analogical examples. However, it is worth noting that most of the tasks used in these studies are primarily analytical and have predetermined solutions. As a result, there is a need for additional research to understand how individuals spontaneously generate or identify connections through analogies in open-ended tasks (Weinberger et al., 2016) and how this process taps into their spatial thinking.

3.2.4 Design ideation through and with analogy

Generating novel design ideas through analogical thinking has nuanced differences from how analogies are used in common problem-solving (Qian & Gero, 1996). In cognitive science and psychology, analogical thinking is often assessed by selecting the correct answer from a range of given solutions (e.g., Simms & Richland, 2019; Yuan et al., 2017) and is, in a way, closed-ended. Designers, on the other hand, appear to go through some distinct cognitive processes when using analogies for design ideation.

In authentic problem-solving contexts, designers typically do not identify or choose from ready-made analogies as would be seen in controlled experiments (Goldschmidt, 2001). Instead, designers usually need to first imagine various possible configurations of the target design before proceeding with mapping between the source and target. The analogy becomes clear in hindsight once a design-by-analogy product is created. However, during the design ideation process, or whenever one is searching for a problem solution mentally, the source and target may not be immediately obvious. Goldschmidt (2001) theorized that the mapping between source and target in design is bidirectional and dynamic, and the details in such structural mapping continue to iterate until finalization (Figure 3.1). Generally, it seems that designers have ample space to explore mapping possibilities before arriving at one solution, such as identifying transferable and designable elements of the source, establishing various levels of mappings, and experimenting with different degrees of adaptation and transformation.

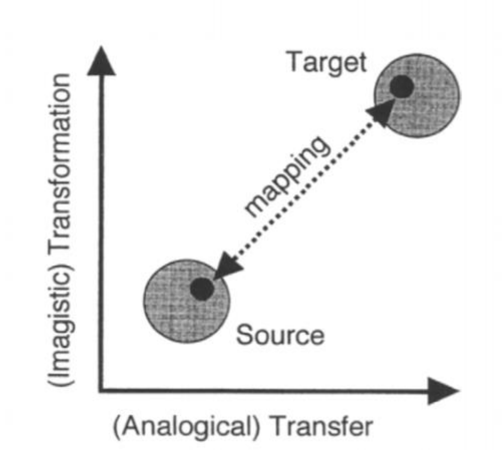


Figure 3.1 Diagram of dynamic imagistic/analogical search processes in creative problem solving (Goldschmidt, 2001, p. 208)

According to Qian and Gero's (1996) model of analogical problem solving and exploring processes in design, designers encode and retrieve functions and features of existing designs as retrieval cues for potential analogical mapping and carryover. Designers then make necessary adaptations or transformations to develop new designs based on the source design. Since design-by-analogy transfers only a certain number of features from the source to the target, designers need careful abstraction (Gentner & Medina, 1998; Qian & Gero, 1996) to determine which features to transform into the new design, disregarding features irrelevant to the analogy. Overall, the exploration of potential analogical mappings and the steps taken to narrow down the final design choice reflect how designers utilize a combination of divergent and convergent thinking processes to draw inferences and transfer insights between domains.

In summary, we have reviewed how spatial thinking plays a role in design ideation as well as in thinking with visual analogies. However, existing research primarily focuses on developing individuals' spatial abilities through convergent, analytical tasks. This leaves a gap in understanding how spatial thinking is used in open-ended problems that require divergent thinking and originate from authentic learning experiences. Furthermore, it appears that the use of analogy, both as a spatial reasoning tool and a design ideation tool, would benefit from further investigation into its usage in authentic and open-ended design problems. To address these gaps, this case study aims to unpack spatial thinking's role in the design ideation process. Specifically, we seek to understand (1) the spatial thinking processes involved in a design-by-analogy ideation process and (2) more broadly, how spatial thinking used in open-ended design tasks differs from that used in tasks with predetermined solutions.

3.3 Current study

3.3.1 Biomimicry as a type of design-by-analogy

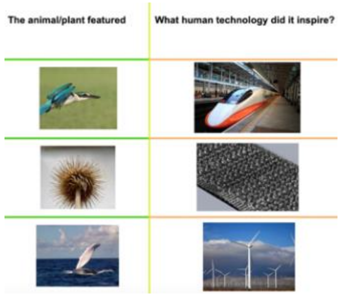
Biomimicry, a form of design-by-analogy, draws inspiration from nature's strategies to devise artificial creations and solutions. It is increasingly recognized as a valuable STEM topic for primary and secondary classrooms (Gencer et al., 2020; NGSS Lead States, 2013). Similar to other analogy-based design processes, biomimicry often begins by examining intricate patterns or structures in organisms in nature. By relating these observed forms to their potential functions, one can then map inferences from nature's form-function relationships onto human designs, resulting in biomimicry designs. This mapping process requires a far transfer of knowledge between nature's strategies and possible human designs, two markedly distinct contexts (Kolodner et al., 2003). To understand the spatial features of organisms in nature, such as their shapes and sizes (Rustaman & Rahmat, 2017), or the three-dimensional structures and processes in biology (Milner-Bolotin & Nashon, 2012), one needs spatial thinking. Engaging students in exploring nature's features for inspiration not only brings their attention to specific features of organisms that they might have overlooked before but also challenges their abilities to draw inferences from and make creative use of nature's form-function relationships.

3.3.2 The design project

We developed this design project following most of the biomimicry design processes listed in the Biomimicry DesignLens, a design guide made by Biomimicry 3.8 (2015). The first and third authors tailored the design project to be suitable for this age group and class structure, resulting in eight 45-min sessions. The design challenge involved asking pupils to create wind- or water-resistant biomimicry designs for camping. Given that the wind or water resistance functions are often performed by perceptible external features of organisms, we anticipate that pupils would use spatial thinking to observe, visualize, and abstract useful inferences from the visual-spatial forms seen in nature. Subsequently, they would apply these inferences to develop designs that incorporate salient visual-spatial features inspired by nature.

To support pupils' understanding of the complex concept of biomimicry and stimulate analogical, divergent, and spatial thinking related to the design challenge, we incorporated a series of scaffolds. Gick and Holyoak (1980) highlighted the importance of explicitly instructing students to reason analogically, as they might not automatically recognize how relevant information can serve as analogous solutions to problems. Other research also endorsed the educational strategy of providing students with ample opportunities to think relationally and visualize analogies across different domains (Stevens et al., 2021, 2022; Vendetti et al., 2015). Therefore, we developed six activities to scaffold pupils' analogical thinking and ideation. Each activity explicitly requires pupils to think analogically (Table 3.1). We expect that these activities will help pupils articulate their intermediate thoughts, gain a better understanding of the core elements of biomimicry designs, establish connections between organisms in nature and the human applications inspired by them, and successfully draw inspiration from nature for their own designs.

Table 3.1 Session details of the biomimicry design project

Session	Learning objectives	Activities that explicitly require analogical thinking	Examples of pupils' work or progress
1	Understanding and examining the design problem; recognizing nature as a source of inspiration for human design	Learning about biomimicry design examples that have been inspired by visible visual-spatial features in nature: burdock burrs inspiring the design of Velcro, Kingfisher's beak inspiring the design of the front of the Shinkansen train, and Humpback Whales' flippers inspiring the redesign of wind turbines	
2	Exploring outdoors and brainstorming for preliminary design ideas	Engaging in a nature walk to observe and note down findings from nature using the "I notice, I wonder, It reminds me of, It inspires me to" worksheet	
3	Investigating organisms in nature for design inspirations; linking biological strategies to design strategies	Analyzing a biomimicry design example in detail to identify the organism, trait, function, biological strategy, and how these elements inspire a design strategy for human design applications	
4	Generating ideas and identifying the best ideas that reflect biomimicry and fit the design challenge	Sketching four different ideas and thinking divergently about different possible mappings from a chosen source	

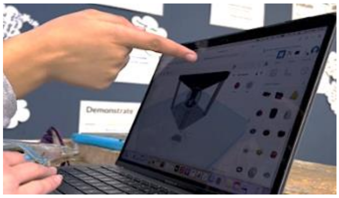
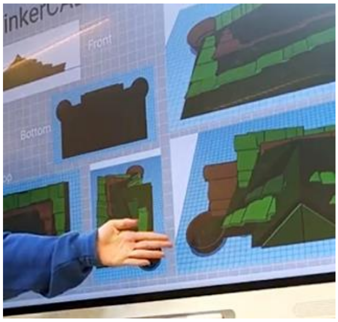
5	Creating detailed sketches and annotations for the selected design ideas; giving feedback on each other's designs	Elaborating on selected design ideas using a design plan worksheet, describing the name, trait, function, and biological strategy of the inspiration source, as well as the feature, function, and design strategy of the design	
6 & 7	Creating and iterating on 3D prototypes of the design ideas using TinkerCAD		
8	Presenting design artifacts in class	Elaborating on design artifacts using a design presentation worksheet, detailing the trait, function, and biological strategy of the source of inspiration, the feature and function of the design, the design strategy, and how the source from nature inspired the design	

Table 3.1 continued

3.4 Methods

3.4.1 Design-based research

We adopted the design-based research approach to translate knowledge from research into educational applications. Design-based research systematically and iteratively tests innovative learning designs, draws from multiple evidence sources in educational contexts to present a holistic view of how learning occurs, and aims to develop actionable and generalizable knowledge regarding educational practices (Bakker, 2018; Design-Based Research Collective, 2003). To ensure the effectiveness and relevance of our design project, we collaborated closely with a Design & Technology teacher to co-create learning materials used in the project and integrate the project into one module within the International Baccalaureate (IB) design course. The learning objectives in this biomimicry design project aligned with IB learning goals, such as learning about forms and functions. Throughout the project, we continuously refined the activity structures based on feedback from the class and the teacher, with the goal of determining how best to support pupils' design processes.

3.4.2 Participants

Sixteen pupils aged 11 to 12 (10 boys, 6 girls) from the first year of secondary school at an international school in the Netherlands consented to participate in this design project. They were joined by their Design & Technology teacher, who delivered all the design sessions. All participants gave researchers permission to use classroom recordings and design works as research data. Most of these pupils had one year of experience in design from their previous school year, which means they were familiar with the design process and certain design techniques, such as sketching and providing design feedback for peers. Pupils worked in duos for this design project, as is typical in design activities. In addition, since many of these pupils had prior exposure to TinkerCAD—a computer-aided design platform where pupilscan combine, resize, and rotate 3D shapes to digitally represent their designs—they were encouraged to build their design prototypes on TinkerCAD.

3.4.3 Data collection and analysis

Data were collected from eight design sessions by the principal researcher and two other researchers. A total of 6 hours and 51 minutes of video and audio data were recorded, with a focus on design activities related to ideation. Classroom notes were taken during the sessions to provide additional context to the recorded data. Pupils' design worksheets, 2D and 3D design artifacts, and self-reflection forms were photographed.

Our analysis centered on pupils' spatial thinking processes in design, specifically on idea generation, elaboration, and development. We conducted an exploratory, data-driven thematic analysis (Boyatzis, 1998), using the qualitative analysis software MAXQDA to analyze classroom videos, audio, notes, pupils' design worksheets, and their 2D & 3D design artifacts. After transcribing the video and audio data, the principal researcher divided the transcriptions into segments. Each segment focused on pupils' discussion on a specific topic before transitioning to the next. Data from other modalities were incorporated into the segments to complement pupils' verbal communications, including pupils' design worksheets with written explanations and sketches, their 2D & 3D design artifacts, and any gestures observed in videos. This multimodal approach aimed to provide a holistic representation of the design processes and pupils' design choices (Van Mechelen, 2016). It is noteworthy that while pupils' spatial thinking can be externally represented through sketches, gestures, or verbal and written communications, a large portion of spatial thinking occurs internally within pupils' minds. Therefore, our analysis addressed data at both semantic (explicit) and latent (interpretative) levels (Boyatzis, 1998; Braun & Clarke, 2006).

The analysis process involved three iterative rounds of coding and discussion between four researchers. The principal researcher identified a range of initial codes related to the design processes as well as to spatial, analogical, and divergent thinking processes. These initial codes were then discussed with three other researchers to identify codes and segments relevant to the research questions. In the second round of coding, the principal researcher reviewed all data segments to refine the code definitions and reduce overlap between codes, resulting in a set of intermediate codes. All four researchers used these intermediate codes to independently code and interpret randomly selected data segments. By comparing the coding results and interpretations, the principal researcher shortlisted five themes that underpinned pupils' spatial thinking during the design-by-analogy ideation process and further refined the definition of these themes.

3.5 Results

All pairs of pupils created design prototypes that demonstrated a proper understanding of designing through biomimicry. That is, instead of merely replicating forms from nature, all pupils created prototypes based on the form-function relationships observed in nature. In their self-reflections, a majority of the pupils perceived their achievements in several areas as medium to high. These included closely observing the characteristics of organisms in nature to uncover new knowledge, applying what they learned from the biological strategies in their designs, generating multiple different design ideas using inspirations from nature, and gaining a fresh perspective on design and technology after exploring biomimicry design examples.

Through the data-driven thematic analysis of pupils' design artifacts and their verbal and written communications, we identified five main themes that capture the meaningful spatial thinking processes underlying the generation and development of their ideas. Two of these themes describe spatial thinking processes specific to analogy-based idea generation, while the remaining three themes describe spatial thinking processes relevant to general design ideation processes. All pupils' names mentioned below are fictitious (Table 3.2).

Table 3.2 Themes of spatial thinking in the biomimicry design project

Theme	Description	Example
Spatial thinking processes pertaining to design-by-analogy		
Abstracting spatial features observed in the source to infer form-function relationships	Observing the forms of the source inspiration and abstracting spatial form-function relationships seen in the source	A duo of pupils noticed that ivies grow in layers and inferred that this form of growth may serve a function of resisting wind.
Developing various possible analogical matches by spatial features and relationships	Exploring different possible analogical matches from a single source of inspiration, or identifying more than one type of form-function relationships that can be analogically applied to inspire design, leading to multiple transformations of mental representations	A duo of pupils generated four different design ideas—tent, camping bag, camping chair, and camping shoes—all inspired by the form-function relationships observed in the spiral-grained growth pattern of pine trees, which allows the tree to be bendable in wind and aids in water distribution.

General spatial thinking processes		
Retrieving relevant visual information from memory	Retrieving a range of visual information from memory that either resembles the desired forms or provides relevant functions	After choosing the pinecone as their source of inspiration, a duo of pupils brainstormed several possible analogical matches based on form or function. They reasoned that while the mechanism of pinecone's scales—closing in wet weather and opening in dry weather—might not be a suitable inspiration for an umbrella design, it could inspire the design of a house or a car that needs to be closed in wet conditions.
Elaborating and iterating on design concepts and representations	Elaborating on and externalizing mental representations of ideas through sketches and annotations, and continuously iterating on these visual representations of ideas	A duo of pupils noticed that the foldings on hornbeam leaves allow the leaves to be flexible in winds and rigid during photosynthesis. Drawing from this observation, they reasoned that a solar panel, serving the same light energy collection function as leaves do, could also be designed with foldings on its surface.
Visualizing the functionality of design prototypes in practical usage scenarios	Visualizing how the design prototypes would be used to evaluate their functionality, consider the user's perspective, and think of possible improvements	A duo of pupils visualized how falling water would interact with the flat-surfaced tarp they designed as well as the trajectory of water streams. They then iterated on the form of their design by elevating the top of the tarp to a pyramid shape and visualized again to consider the functionality of their improvement.

Table 3.2 continued

3.5.1 Spatial thinking processes pertaining to design-by-analogy

Abstracting spatial features observed in the source to infer form-function relationships

Throughout this design project, pupils actively observed various forms in nature and abstracted the spatial form-function relationships from these sources of inspiration. After returning from their outdoor exploration in the second session, all the pupils sketched and annotated the inspirations they gathered from nature. Below is an example of a child's observation and reasoning of ivy leaves and long grasses (Figure 3.2).

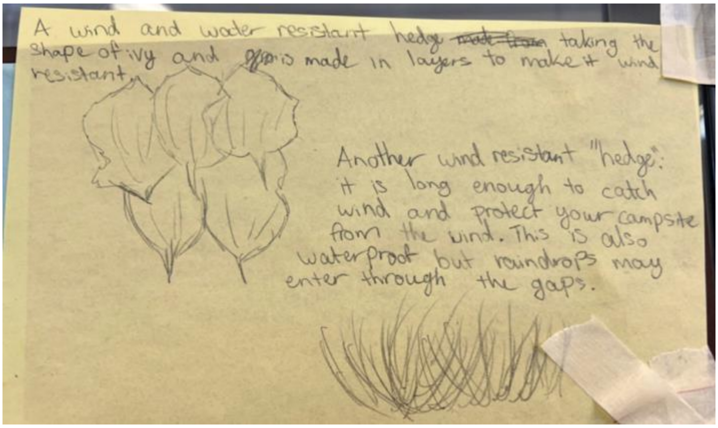
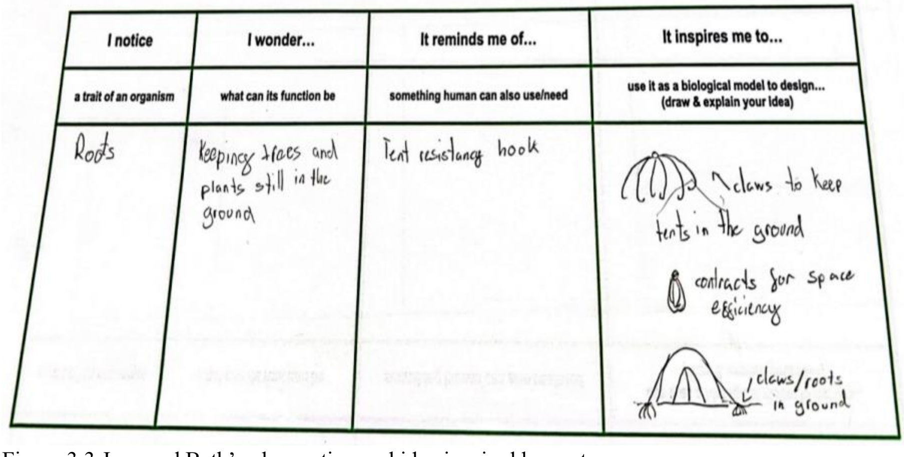


Figure 3.2 Iris' observation of ivy leaves and long grasses

Out of the many characteristics of these plants, such as color and texture, Iris identified the spatial features that ivies grew "in layers," and the dense hedge of grasses was "long enough to catch wind." Based on these observations, she reasoned that these features might serve functions related to wind and water resistance. The descriptions of layers of growth and the wind-catching feature further suggest that she may have observed or visualized in her mind how wind dynamically interacts with these plant hedges. These keen observations paved the way for further exploration of how forms observed in nature can potentially serve wind- or water-resistant functions. In this particular example, Iris theorized that the long grass hedge could function as a protective barrier for a campsite. Although she may not have fully grasped the exact mechanisms behind the wind and water resistance of these plants, her sketches and descriptions revealed how she was thinking spatially about the shapes and structures in nature, as well as using these abstracted form-function relationships from nature in potential design applications.

Jean and Beth directed their attention to the shape of the roots. They recognized how the shape of roots helps stabilize plants in the ground against wind and reasoned that it serves the function of "keeping trees and plants still in the ground" (Figure 3.3).



I notice	I wonder...	It reminds me of...	It inspires me to...
a trait of an organism	what can its function be	something human can also use/need	use it as a biological model to design... (draw & explain your idea)
Roots	Keeping trees and plants still in the ground	Tent resistancy hook	claws to keep tents in the ground; contracts for space efficiency; claws/roots in ground

Figure 3.3 Jean and Beth's observation and idea inspired by roots

Guided by the prompts in the worksheet, they associated the form-function relationship of tree roots with that of a hook, which is commonly used for stabilization, such as in tents. Interestingly, their sketch of the designed tent hook resembled the form of plant roots. They used sketches and annotations to visually elaborate on how the design would be used, detailing the steps of the hook being inserted into the ground to keep the tent stable, as well as the hook's ability to contract for space efficiency when not in use. These pupils' analogical exploration of forms and functions and their abstraction of form-function relationships indicated that they thought spatially about the form, position, and usage of the source (tree roots) and the target (tent hooks).

It is noteworthy that all pupils in the design project sketched or described new insights they learned about the features of organisms in nature. These examples highlight that developing design-by-analogy ideas typically requires understanding the visual-spatial features of the source, which, in this case, are organisms in nature. Close observations of the forms of these sources of inspiration and meaningful abstractions of spatial form-function relationships laid the groundwork for exploring potential analogical mappings and design ideas.

Developing various possible analogical matches by spatial features and relationships

After familiarizing themselves with the design problem and seeking inspiration from nature, the pupils embarked on a search for potential design solutions by exploring various possible analogical matches based on spatial features and relations. As an example, a duo of pupils generated four different design ideas inspired by the spiral-grained trunk of the Whitebark Pine (Figure 3.4).

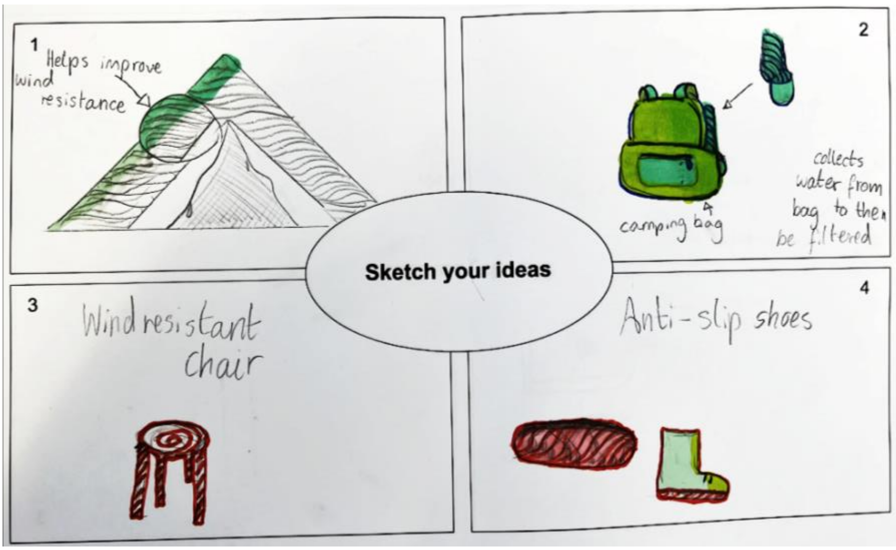


Figure 3.4 Four different design ideas inspired by the spiral-grained trunk of Whitebark pine by Jo and Sanne

It is important to note that, while Jo and Sanne's inspiration stemmed from the spiral-grained tree trunk, their design ideas did not solely replicate its rigid shape. Having learned that the spiral form on the tree trunk allows for more bending and better distribution of pressure when faced with strong wind, they explored various ways to embody similar form-function relationships in their designs. They envisioned non-rigid spatial transformations of the spiral pattern and applied them to diverse contexts, such as the surface of a pyramid-shaped tent, part of a camping bag, and the soles of shoes. By adapting these biological strategies into design strategies, they reasoned that a tent with an uneven, spiral-grained surface might enhance wind resistance, and a chair with spiral-grained legs could offer more stability in camping scenarios.

Interestingly, ideas 2 and 4 (see Figure 3.4) were derived from two additional form-function relationships they identified. Jo and Sanne learned that the spiral-grained tree trunk facilitates the distribution of water across the trunk and its branches. Drawing on this observation, they developed idea 2, which incorporated a spiral-grained water-collecting and filtering device into a camping bag design. As for idea 4, these pupils speculated that the spiral pattern may enhance traction, even though the information about spiral-grained pine trees did not explicitly mention grip. Based on this hypothesis, they developed idea 4, where the spiral pattern was utilized to improve the shoe's grip.

This example highlights these pupils' ability to make use of more than one type of form-function relationship, exploring diverse ways to integrate spatial features into their own designs. By engaging in an open-ended search for various solutions, these pupils were able to visualize numerous spatial transformations in their minds.

3.5.2 General spatial thinking processes

The two themes discussed above reflect spatial thinking related to the design-by-analogy process. We will proceed to discuss three themes that were important to this specific design challenge but also appear to be applicable to many other design ideation processes.

Retrieving relevant visual information from memory

After being introduced to the design challenge, the pupils retrieved relevant information from their memory that was visually related to the desired form or matched the description of relevant functions. The example below shows a pair of pupils contemplating design ideas inspired by the pine cone, which has scales that open in dry weather and close in wet weather.

Hugh: look it cannot be an umbrella, because when it's cold and wet it's gonna close
Ellen: a house, okay never mind, oh a car, like a folding car, it opens when it's warm and closes when it's cold and rainy… like a convertible car…

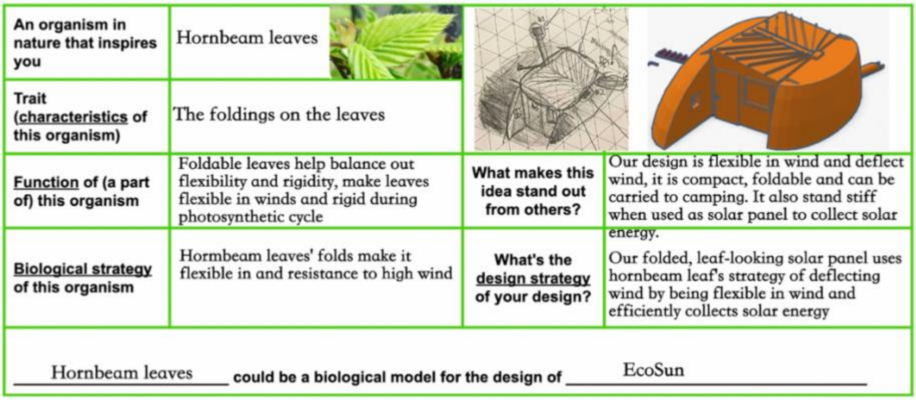
From their conversation, we can infer that these pupils had a basic mental representation of pine cones' biological strategies in their minds. They retrieved various visual information from their memory, such as umbrellas, houses, and cars, to see if the design strategies of these items aligned with pine cones' biological strategies. During their discussion, Hugh quickly noticed that umbrellas open and close in a way opposite to the scales of pine cones. While the idea of a house seemed feasible, Ellen appeared to be more enthusiastic about the idea of a car. He reiterated the mechanism of how the car would open and close, possibly to confirm the compatibility of the two mechanisms.

Through this process of considering how the pine cone's biological strategies could be applied to their designs, Hugh and Ellen compared their mental representations of the form-function relationships observed in the source (pine cone) with potential targets (umbrella, house, car). They engaged in mental imagery to explore possible design ideas, likely visualizing the opening and closing mechanisms, imagining possible motions, and envisioning scenarios where these spatial changes would occur. Based on this comparison and reasoning, they quickly concluded that the umbrella was not a suitable fit. Given the open-ended nature of this design prompt, they further explored the house and the car ideas, as both offered potential for valid analogical mappings.

By tapping into their stored visual knowledge, these pupils drew from existing patterns and forms to inspire their design solutions. While closed-ended spatial tests may also require retrieving relevant visual information from memory, the process of thinking divergently about many possible solutions encouraged the pupils to conduct a broader search in their memory for relevant visual information.

Elaborating and iterating on design concepts and representations

As the pupils explored different design options and ultimately decided to further develop one idea, they adapted inferences drawn from the form-function relationships observed in nature to their designs and engaged in a deeper level of reasoning within their design-by-analogy process. Below is a design plan worksheet from a duo of pupils who designed a foldable solar panel inspired by hornbeam leaves (Figure 3.5).



An organism in nature that inspires you	Hornbeam leaves		
Trait (characteristics of this organism)	The foldings on the leaves		
Function of (a part of) this organism	Foldable leaves help balance out flexibility and rigidity, make leaves flexible in winds and rigid during photosynthetic cycle	What makes this idea stand out from others?	Our design is flexible in wind and deflect wind, it is compact, foldable and can be carried to camping. It also stand stiff when used as solar panel to collect solar energy.
Biological strategy of this organism	Hormbeam leaves' folds make it flexible in and resistance to high wind	What's the design strategy of your design?	Our folded, leaf-looking solar panel uses hornbeam leaf's strategy of deflecting wind by being flexible in wind and efficiently collects solar energy
Hornbeam leaves could be a biological model for the design of EcoSun			

Figure 3.5 Design plan for the foldable solar panel inspired by hornbeam leaves, created by Sean and Andi

Sean and Andi demonstrated a clear understanding of the form-function relationships in their source of inspiration. They recognized that the foldings on hornbeam leaves provide flexibility in wind and rigidity during photosynthesis. This understanding encompassed the spatial features of the source, the function denoted by the folding structure, and a basic mental representation of wind dynamically interacting with the leaves. As these pupils elaborated on the biological strategies and design strategies, they explicitly mentioned that just as the foldings on hornbeam leaves make the leaves flexible in the wind and rigid during photosynthesis, their solar panel design, which serves a similar function to leaves in collecting light energy, can also benefit from having foldings on its surface. They further reasoned that these foldings offered additional benefits to the solar panel, allowing for a foldable and compact design that is easily portable for camping.

Their elaboration on the design plan indicates that Sean and Andi were not only thinking spatially about the form of their design but also tailoring their design concept to the context of this design challenge.

Another example of elaboration and development of the design concept can be seen in the progression in sketching (Figure 3.6). Jo and Sanne's sketches evolved from an initial concept to a more detailed visual representation. They not only added three-dimensionality to their sketch but also incorporated additional features to the tent, such as a clear entrance, attached strings, and buckets for water collection. Their understanding of the spiral patterns on the tree trunk, which aid in water and nutrient distribution, influenced the development of their design idea. Both Figures 3.5 and 3.6 highlight how these pupils not only externalized the mental representations via sketches but also further expanded on their design ideas with increasingly detailed visualizations.

Source | Initial Sketch | Final Sketch

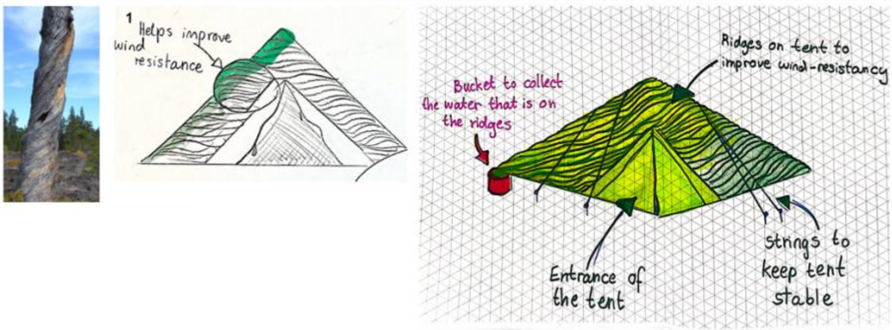


Figure 3.6 Development of ideas seen from Jo and Sanne's design sketches

Visualizing the functionality of design prototypes in practical usage scenarios

According to Finke et al. (1992), the creative thinking process includes envisioning scenarios in which the functionalities of a new design are explored. In our data, we also observed pupils visualizing their design prototypes in use, evaluating functionality, considering the user's perspective, and thinking of possible improvements. Below, we present a sketch, a TinkerCAD model (Figure 3.7), and a conversation segment from a duo of pupils who designed a camping tarp inspired by the spiral-grained trunk of Whitebark pine. They based their design on the observations that the spiral-grained tree trunk offers wind resistance and facilitates water distribution.

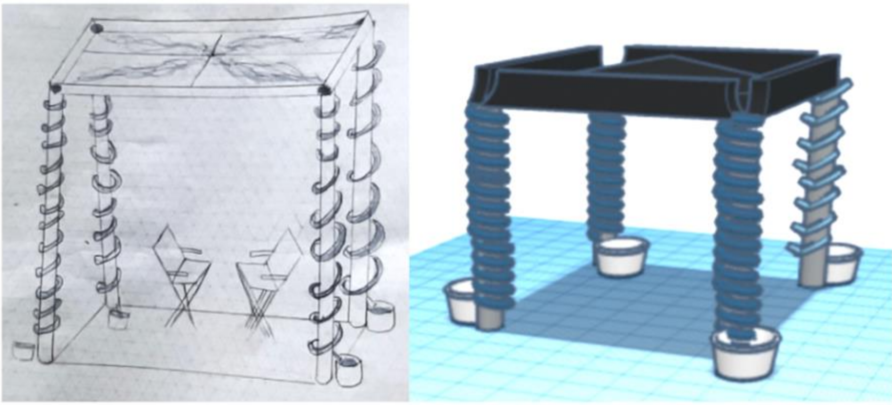


Figure 3.7 Sketch and TinkerCAD artifacts from Beth and Ivy, showcasing their camping tarp design inspired by the spiral-grained trunk of Whitebark pine

On TinkerCAD, Ivy was resizing the pyramid-shaped top, so it became larger and its four angles reached the four corners of the box.

Beth: does it make sense?... No you know what it doesn't work? because all the water will go over there, over there, and over there, and over there, and not to the corners
As they spoke, Beth pointed at the screen to gesture how the falling rainwater would go to the four sides of the pyramid, horizontally rotating the shape on the top
Beth: now it's perfect you see (dragging the design to check from the side and the top) so the water always goes into that direction, that direction, that direction and that direction...

From their conversation, it is clear that as Beth and Ivy progressed from the sketch on the left to the TinkerCAD design on the right (Figure 3.7), they modified the camping tarp's top design into an elevated pyramid shape. These verbal, virtual, and bodily representations show that Beth was actively visualizing the scenario where the design would be used. She visualized how rainfall would interact with the top of the tarp, as well as the trajectory of water flow, concluding that an elevated pyramid-shaped top would address the issue of water overflow. This practice of validation through visualization not only engaged pupils' spatial thinking but also helped them identify potential areas of improvement in their sketch and TinkerCAD model. Subsequently, it prompted them to iterate on the form of their design to ensure its functionality and cater to users' needs.

3.6 Discussion and conclusion

This case study aims to bring together knowledge from the field of spatial ability research, creative cognition, and analogical reasoning. Through a triangulation of evidence from our data, we observed that design education offers unique and largely untapped opportunities to foster spatial thinking development even from an early age, given the social-material nature of design. Pupils' development of ideas through design-by-analogy demonstrated their frequent use of spatial thinking. They closely examined the spatial features of organisms in nature, abstracted the implications of these spatial forms for specific functions, and visualized—both in their mind and through sketches or modeling—how these form-function relationships could be transformed and applied to a variety of design ideas. Furthermore, the detailed visualizations of ideas created by these pupils (Figures 3.2—3.7) resonate with Clement's (2008, 2009) proposition that imagistic mental simulations and imagistic spatial transformation are crucial for generating analogies to solve problems creatively. Meanwhile, we also noticed that the types, processes, and goals of spatial thinking seen in open-ended design ideation have nuanced differences from those in closed-ended spatial tasks.

The first theme, abstracting spatial features observed in the source to infer form-function relationships, lays the foundation for analogical mappings. A close examination and abstraction of the key visual-spatial features or processes from biological examples is essential for understanding biological strategies and adapting them into design strategies (Stevens et al., 2021, 2022). Meanwhile, this theme also aligns with previous findings suggesting that design-by-analogy transfers only a certain number of features from the source to the target, requiring careful abstraction from designers (Gentner & Medina, 1998; Qian & Gero, 1996). To solve spatial tasks, observing and abstracting visual-spatial features is important, as one may derive the correct answer to a spatial problem by visualizing the transformation of a key feature (Barratt, 1953). However, in design-by-analogy, pupils observe and abstract spatial features with the unique goal of discerning their functional roles in nature, and subsequently transforming such form-function relationships into meaningful new designs. This is akin to what Finke et al. (1992) described as mentally re-arranging and transforming what was known into something new. Hence, it appears that the objective of observations and abstractions in solving open-ended design tasks is distinct from that in conventional, closed-ended spatial tasks.

In contrast to the use of spatial thinking in closed-ended spatial tests, the themes, retrieving relevant visual information from memory and developing various possible analogical matches by spatial features and relationships, suggest that the search for possible solutions in open-ended tasks likely requires pupils to recall and browse through a wide array of relevant visual information. During this process, pupils might need to form, retain, and manipulate multiple mental representations—which act as 'temporary spatial displays' (Kosslyn et al., 1979)—before deciding on and finalizing the desired idea. By contrast, conventional spatial tasks typically present a selection of given choices. As Maresch (2014) reviewed and summarized, individuals can employ various strategies to tackle a spatial task, including using a falsifying strategy to rule out impossible answers or prioritizing the verification process once a potentially correct answer is identified, thereby dedicating less attention to other options. While these strategies may be effective for multiple-choice questions, they might not be applicable to open-ended prompts. Therefore, we argue that existing spatial training and tasks might not sufficiently prepare pupils for the visually and spatially demanding processes in design ideation. Such ideation often necessitates a broad exploration of visuals in different directions (Gabora, 2010) and is essential for creativity and innovation.

To share products of creative visualization with others, designers need to "mentally create, manipulate and communicate solutions effectively" (Isham, 1997, p. 2). Similarly, as these pupils developed their ideas, they continued to enrich the elaborations of their selected ideas by refining details, evaluating suitability, and brainstorming possibilities for improvements and iterations. The themes, elaborating and iterating on the design concept and representations and visualizing the functionality of design prototypes in practical usage scenarios, engage both divergent and convergent thinking. Prior research (Cropley, 2006; Goldschmidt, 2016; Schut et al., 2020) indicated that convergent thinking plays an important role in the design process for generating viable ideas. Yet, we observed an interesting distinction between the convergent thinking applied in design ideation and that in conventional spatial tasks. In spatial tasks, convergent thinking is frequently used to determine a single correct answer, typically without the need to refine or improve on the answer. In design, pupils converge on their design concepts and visualize the functionality with the goal of identifying areas of improvement. This often leads to thinking spatially about possible improvements in their designs' forms (e.g., Figures 3.6 and 3.7). Once again, existing spatial training and tests may not adequately reflect how pupils employ spatial thinking to assess their design ideas or how they actively consider potential improvements to the visual-spatial features of their ideation products.

While spatial thinking is often trained with the goal of enabling learners to perform all transformations mentally, we must not overlook the important role embodied experiences play in spatial thinking (Frick et al., 2009; Link et al., 2013). In this design project, some design ideas stemmed from pupils' embodied interactions with organisms in nature. This leads us to wonder how these embodied interactions might have supported pupils' spatial thinking, as well as their analogical and divergent thinking. As Pallasmaa (2017) described, "In the arduous processes of designing, the hand often takes the lead in probing for a vision, or a vague inkling, which it eventually turns into a sketch, materializing thus the idea" (p. 104). Therefore, instead of treating bodily motions and cognitive capacities as distinct agencies, further investigation is needed, perhaps through a multimodal approach, to understand how embodied experiences in design may support pupils' spatial thinking processes.

In summary, this case study delved into the spatial thinking processes involved in a design-by-analogy ideation process and identified differences between spatial thinking observed in open-ended design ideation and that typically assessed or trained in existing spatial ability research. The potential differences discussed above warrant attention from future researchers, especially when assessing spatial ability or developing spatial training, given that many real-life problems are open-ended, requiring a combination of divergent and convergent thinking. Presently, our understanding of spatial thinking largely derives from psychometric test results. By depicting these themes of spatial thinking observed in design ideation—themes not previously discussed in traditional spatial ability research—we suggest the potential need for a more comprehensive definition of spatial thinking. This might be especially relevant for the field of design or other fields that rely heavily on divergent thinking. Admittedly, our proposed themes might not be exhaustive, as our investigation focused primarily on design ideation. Future research is needed to uncover other spatial thinking processes seen in design iteration or design feedback phases.

Lastly, although the concept of biomimicry design was new to the pupils who participated in this design project, our case study suggests that pupils of this age can develop biomimicry designs when appropriate scaffolding is provided throughout the project. One type of scaffold came from the materials developed for this project. For instance, the worksheets we designed emphasized the form-function relationship and the analogical mapping between biological strategies and design strategies (e.g., Figures 3.3 and 3.5). Another type of scaffolding involved utilizing teacher's support to enhance pupils' learning throughout the design process. In this specific design challenge, the teacher's language and questions directed pupils' focus on the visual-spatial forms, such as spirals, folds, and layers of scales, and how these forms imply water- or wind-resistant functions. Such guidance ensured that the pupils focused on the key form-function relationships without being distracted by other characteristics of the organisms that were less pertinent to the design challenge. The teacher's support also stimulated valuable realizations and conversations among the pupils. For instance, to help a pair of pupils understand that biomimicry does not mean directly using organisms in nature (e.g. using pine cones as weather-tellers) but rather applying what can be learned from pine cone scales' weather-responsive mechanism to create a new design, the teacher reminded them of a prior design-by-analogy example they had learned about, in which the human femur bone and the Eiffel Tower may share a similar efficient structure but are made of distinct materials in different sizes. Subsequent discussions between these pupils reflected a clearer understanding of biomimicry. Practices like these align with the recommended strategies to support analogical learning (Vendetti et al., 2015). It is necessary to note that our analysis focused on the progress and artifacts produced by pairs of pupils rather than individual efforts. Thus, future research is needed to investigate how to scaffold individual children's spatial thinking, divergent thinking, and analogical thinking within a design project.

CHAPTER 4
Evaluating pupils’ creative design ideation and prototypes through comparative judgement

CHAPTER 4

Evaluating pupils' creative design ideation and prototypes through comparative judgement[4]

Developing creative solutions to improve our surroundings is a key 21st-century competency. Design & Technology (D&T) presents many opportunities to teach creativity as a skill. Yet the ill-defined and context-dependent design problems often make it challenging for educators to adequately evaluate the creativity seen in pupils' solutions. The method of comparative judgement does not require a predetermined set of evaluative criteria. In this study, we leveraged this technique to explore the evaluative considerations that 20 expert judges used in their holistic assessment of design ideas and prototypes produced by 201 pupils (M_{age} = 12.52) in the Netherlands. Despite creativity's recognized importance in shaping design quality, it is not prioritized in many current D&T projects. Therefore, we deliberately focused on evaluating the creativity seen in these pupils' works. We further explored how judges' consideration, coded as criteria, shifted from the beginning to the end of the comparative judgement process. Our findings from qualitative and quantitative analyses together added to our understanding of the multifaceted process of evaluating creativity and provided practical insights into using comparative judgement as an assessment method in design education.

[4] *This chapter includes text that has previously appeared in the following source:*

Zhu, C., Buckley, J., Klapwijk, R., Spandaw, J., de Vries, M. J. (2025). A holistic look at creativity: Assessing pupils' creative design ideation and prototypes through comparative judgment. International Journal of Technology and Design Education.

4.1 Introduction

Creativity—one of the key 21st-century competencies (Voogt & Roblin, 2012)—is widely recognized as a crucial factor driving innovative designs (Christiaans, 2002; Cropley & Cropley, 2010; Dorst & Cross, 2001; Kimbell & Stables, 2007; Lewis, 2005; Sarkar & Chakrabarti, 2011). The National Curriculum for Design and Technology subject in England, for instance, highlighted one of its learning goals as for pupils to "develop the creative, technical and practical expertise needed to perform everyday tasks confidently and to participate successfully in an increasingly technological world" (Department for Education, 2013). Design and Technology (D&T) classrooms offer untapped potential for exploring how pupils utilize and develop creativity (Benson & Lunt, 2011; Cropley & Cropley, 2010; Lewis, 2005, 2009; Xu et al., 2020). However, adequately evaluating creativity in design projects remains a challenge to this day, both in selecting the appropriate task for assessment and in determining how to weigh the various outputs generated from these projects (Casakin et al., 2010; Kimbell & Stables, 2007).

To assess creativity as a cognitive or psychological construct, psychologists have developed test-based measures that are context-independent, making them generalizable across large samples with established scoring guidelines that can be applied by trained raters. However, these creativity measures may not fully capture the nature of creativity within the D&T learning environment, where projects are not only open-ended but also context-specific and rooted in authentic real-world problems (Kimbell & Stables, 2007). To some extent, the design task itself influences how creative a design product can be (Cropley, 2005). Design tasks having a few abstract constraints (e.g., easy to use) seem to have elicited more creative designs than those imposing more concrete constraints (e.g., budget) (Starkey et al., 2016). Selecting or developing a task to assess design capabilities—and choosing the appropriate method for evaluating a complex construct like creativity, whether it applies to the product, the designer, or the overall design process—requires careful consideration that takes into account the progress made over the past two decades in the D&T community.

4.1.1 Assessing qualities of work in D&T projects through comparative judgement

Evaluating responses to open-ended tasks, especially those that require creativity or divergent thinking, has been a challenging topic for both educators and researchers (Jones & Alcock, 2014; Lesterhuis et al., 2017). Traditional methods that rely on predefined scoring rubrics can often be time-consuming and costly, accompanied by complications that arise from large disagreements among judges and judge fatigue (Bejar, 2012; Forthmann et al., 2017). Evaluators may be inconsistent with how they interpret the rating scale throughout the assessment (Hoskens & Wilson, 2001; Myford & Wolfe, 2003; Pollitt, 2012); or they may choose to rely on more straightforward and intuitive criteria, such as aesthetics or crafting quality, since characterizing constructs like creativity can be more complex and intricate (Casakin et al., 2010). Assessing designs by calculating the relative frequency of attributes, such as the relative rarity of a solution, appears to be more objective (Shah et al., 2003). However, these values need to be interpreted with caution or may otherwise lead to a biased assessment of the construct (Sluis-Thiescheffer et al., 2016).

In contrast to evaluations based on predefined criteria, comparative judgement leverages judges' expertise to holistically select the 'better' piece of work through rounds of paired comparisons, thereby revealing the relative quality of works through rank order

(Hartell & Buckley, 2021). Comparative judgement validly and reliably assesses aspects that judges consider central to the evaluated construct (Pollitt, 2012) and appears to be well-suited for constructs that can be hard to precisely characterize, such as creativity (Jones & Alcock, 2014). Studies employing adaptive comparative judgement—a form of comparative judgement with an adaptive algorithm managing the creation of pairs of work and their presentation to assessors—to evaluate the qualities of works from design and technology education in secondary or higher education contexts have reported high inter-rater reliability, ranging from 0.93 to 0.97 (Bartholomew et al., 2018; Bartholomew et al., 2019; Buckley et al., 2022; Seery et al., 2019).

Most studies in D&T education that use comparative judgement have focused on evaluating the overall quality of design works (e.g., Kimbell, 2012; Seery et al., 2019; Strimel et al., 2021). Since many factors can influence judges' perceptions of quality, relying on judges' expertise for holistic assessments implies that they may apply diverse criteria during evaluation (Jones & Alcock, 2014; Lesterhuis et al., 2022). Qualitative findings from Buckley and colleagues (2022) revealed the diverse criteria judges used to select a winning design portfolio from a pair. These included the quality of crafting seen in the work, the quality of the design concept, effectiveness in communication and presentation, the emotion conveyed through the design, and the amount of effort or details evident in the design. Interestingly, how creative, unique, interesting, or adventurous the designs were seemed to be less frequently mentioned as a criterion. Similarly, in the qualitative analysis reported by Bartholomew and colleagues (2018), it appeared that more emphasis was placed on the suitability and feasibility of the design, the aesthetics, and the completeness of the portfolio, with creativity and innovation comprising as little as 5% of the comments made by judges. Another study revealed cultural differences in the design values identified by judges from different backgrounds, reflecting the varying perspectives on what constitutes quality design (Bartholomew et al., 2020). For instance, while D&T experts from the U.K. frequently emphasized innovation in their adaptive comparative judgement process, experts from Sweden and the U.S. mentioned innovation less often, instead prioritizing usability, adherence to design criteria, and effective idea communication.

These findings are somewhat surprising, given that creativity is widely recognized as a vital, even central, component of design quality (Christiaans, 2002; Goldschmidt & Tatsa, 2005; Kimbell & Stables, 2007; Sarkar & Chakrabarti, 2011). Lewis (2005, 2009) emphasized that while creativity is central to design and technology, it remains insufficiently addressed in teaching and learning, and thus demands more explicit curricular and pedagogical attention. Cropley and Cropley (2010) further highlighted assessing the creativity of design solutions as a key strategy for fostering creativity in technological design education. As recent comparative judgement studies (e.g., Bartholomew et al., 2018; Buckley et al., 2022) revealed that creativity-related criteria accounted for only a small portion of design quality evaluation, the question of what shapes our perception of creativity in today's D&T projects remains to be answered. To address this gap and gain deeper insight into how creativity is expressed in young pupils' design work, we proposed using comparative judgement specifically to assess creativity in design. This focus is novel, as comparative judgement studies in D&T education have primarily concentrated on overall design quality rather than creativity as a distinct construct. Furthermore, as comparative judgement reflects judges' conceptualizations of complex constructs (Lesterhuis et al., 2022) and has been successfully applied to assess solutions to various types of open-ended problems (Kimbell, 2012; Steedle & Ferrara, 2016; Strimel et al., 2021), this method can be particularly well suited to unpacking the multifaceted nature of creativity in design.

This investigation is part of a larger research project examining the link between creativity in design and spatial ability. Sections regarding the selection, adaptation, and implementation of the design task were previously presented at a conference (Zhu & Klapwijk, 2024). In this article, we first examined the evaluative considerations used by judges when applying comparative judgement to assess creativity in pupils' design ideation and prototypes. Design ideation refers to the generation of preliminary concepts that may be further developed or discarded in subsequent stages of the design process (Lawson, 2005). Prototyping takes ideas one step further by embodying their role, implementation, as well as look-and-feel in visual or tangible forms (Houde & Hill, 1997). The focus on these early stages of design was because quality ideas are foundational to creative design outcomes (Goldschmidt & Tasta, 2005), and prototypes serve a range of meaningful roles throughout the innovation process (BenMahmoud-Jouini & Midler, 2020). We then examined how judges' evaluative considerations—coded as criteria—evolved from the beginning to the end of the comparative judgement process. Specifically, we investigated whether emphasizing creativity as the main assessment goal allowed judges to evaluate this construct through comparative judgement validly. In addition, we explored whether comparative judgement analysis may offer complementary insights to traditional grading methods by capturing nuanced differences in design ratings, enriching our understanding of design creativity, and revealing the unique challenges judges encounter during the comparative judgement process.

4.2 Methods

4.2.1 Participants and setting

The design project was conducted in four international schools in the Netherlands. A total of 201 pupils (M_{age} = 12.52, SD_{age} = 1.14, 108 girls, 93 boys) completed all parts of a design task and other relevant tasks. These pupils came from 33 different nationalities, and 46.3% of them reported English as one of their native languages. On average, they had been enrolled in one or more international schools where English was the main language of instruction for 5.18 years (*SD* = 2.80). All design activities were conducted in English, and comprehension of the design brief and task prompts was essential for meaningful participation. While pupils with limited English proficiency were supported by classroom teachers and allowed to take part, their data were excluded from the analysis.

Twenty master's students from the Department of Industrial Design Engineering at Delft University of Technology were recruited as judges. Participation opportunities were advertised within the department and via online student groups. Design students who expressed interest were invited to a brief interview and a 30-minute training session led by the research team. This session included an overview of the design assignment and evaluation objectives, along with five practice trials using the comparative judgement platform to ensure that they understood the evaluation process and developed a basic familiarity with the pupils' design levels. Judges were selected based on their demonstrated understanding of the comparative judgement procedure. Of the final panel of 20 judges, five had recently completed their master's degrees, while the remaining were current master's students. Thirteen were enrolled in the Design for Interaction track, five in Strategic Product Design, and two in Integrated Product Design. Twelve judges had prior experience with coursework or design projects related to designing for or with school-age pupils. All judges received hourly compensation for their participation.

4.2.2 Selecting and adapting the design task

The SnackSafe on the Beach design task used in this study was inspired by a real-world problem—annoyances caused by seagulls. This authentic and engaging design challenge was developed by education researchers at the Science Hub at Delft University of Technology in collaboration with Ontwerpen in de Klas, originally instructed in Dutch, and has been carried out in multiple Dutch primary and secondary schools. During the design task selection process, we piloted this activity alongside other potential design tasks. Our aim was to identify a task that could stimulate creative design thinking by encouraging diverse interpretations and a variety of problem-solving strategies, thereby fostering divergent thinking (Klapwijk et al., 2021). Several alternative tasks were excluded because they imposed complex constraints that were difficult for pupils to address within the allotted time, or they tended to elicit similar types of solutions, thereby limiting creative variation.

For research purposes, we made several adaptations to the seagull annoyance design brief and procedure. First, we focused the design task on helping people enjoy their fries on the beach without being bothered by seagulls (Figure 4.1). This specific design problem has previously been tested by educators at the Science Hub and was found to be personally relevant to pupils, feasible to complete individually at their age level, and effective in eliciting a wide range of creative responses. Second, we identified key steps in problem exploration, idea generation, and concept development that could be done individually. These steps were incorporated into an A3-size foldable design booklet to capture pupils' thought processes when solving the design task. To assess individual pupils' performance, each pupil completed all steps independently without collaboration. The design of the booklet was inspired by the design portfolio developed in the TERU project (Kimbell & Stables, 2007; Stables & Kimbell, 2000), which aimed to document pupils' idea progression, intermediate products, and reflections to illustrate the dynamic design process. A copy of the design booklet used in this study can be found in the Appendices.

4.2.3 Organizing the design task and the comparative judging sessions

The energizer and the design task

All design sessions began with a five-minute energizer activity intended to stimulate pupils' thinking and engagement. This practice aimed to help pupils feel comfortable expressing ideas and actively thinking in different directions (Klapwijk et al., 2021). Pupils were asked to list as many fruit names as they could think of within one minute and stand up when the time was up. Those who had either 'apple', 'banana', or 'strawberry' among their first three responses were then asked to sit down. Following this activity, pupils were told: "Oftentimes when we are asked to come up with many ideas, our first few ideas are not very uncommon or unique. However, by continuing to generate more ideas, you may come up with more unusual, uncommon, and creative ideas, or you might be able to make links between ideas to create unique combinations."

Next, the pupils were guided through the design booklet. The researcher read the design problem aloud and instructed them to spend three minutes writing down keywords for the features they believed their design should include to address the design problem.

Figure 4.1 The SnackSafe on the Beach design task

Pupils were then asked to brainstorm as many ideas as possible on an A4 sheet with six empty boxes. They were encouraged with the prompt: "All ideas are welcome. Don't hesitate to sketch anything that comes to your mind. Every new idea can be a valuable addition." This prompt drew inspiration from known brainstorming rules advised by experienced designers, design educators, and researchers (IDEO.org, 2015; Klapwijk et al., 2021).

Following the initial brainstorming, pupils were explicitly instructed to be creative and generate four additional ideas, which could either be entirely new concepts or improvements on previous ideas. Pupils were further advised: "Try to think of unusual, original, exciting ideas that others won't easily think of. Think from different viewpoints and explore different directions. You can also jot down improved versions of your previous ideas or make unique combinations between your ideas." This instruction, adapted from Butler (1987) and drawn from the YourTurn Design Tool, developed through a collaboration between Delft University of Technology and Goldsmiths, University of London (Klapwijk et al., 2021), aimed to clarify the task objective, sustain engagement, encourage divergent thinking, and emphasize the importance of exploring different directions while postponing judgement.

After generating ideas, pupils were instructed to individually select the idea they considered the most original in addressing the design task. They then created detailed sketches of their chosen ideas, adding annotations to help others understand the features and intended functions of their design prototypes. Each pupil had 40 minutes to complete both ideation and prototype sketching, using only pencils for all drawings and notes. All pupils received equivalent instructions, materials, and resources. The above-mentioned prompts were intended to facilitate their idea development and reflect typical practices in design projects.

The comparative judging process

All ideas and design prototypes were photographed, anonymized, and uploaded to an online comparative judgement platform, No More Marking (n.d.), to create two comparative judgement sessions. When pupils' handwritten annotations were hard to distinguish, we transcribed them and placed the typed text alongside the original handwriting for clarity. Figures 4.2 and 4.3 illustrate the interfaces presented to the judges, where pupils' designs were randomly paired across multiple rounds for judges to determine the more creative one in each pair.

Judges were told to use their own understanding of creativity—which can be criteria developed from their personal or professional experiences—to evaluate the designs. The prompts displayed above each pair of works—either 'Which set of brainstormed ideas do you think is more creative?' or 'Which design prototype do you think is more creative?'—were made to maintain a holistic focus on assessing creativity.

Each judge completed 101 pairwise comparisons ([201 pieces of work × 10]/20 judges) separately for pupils' ideation and design prototypes. Judges were required to leave comments on the first and final 15 judgements they made to explain the criteria they used in making their decisions. They were also encouraged to leave additional comments on other comparisons if they felt that it would help the researchers better understand their decision-making rationale. All judges first completed the judging session on design prototypes, followed by the session on design ideation (brainstormed ideas). This sequence was intentional. During the ideation phase, pupils were not required to provide annotations and were encouraged to produce only simple sketches. As a result, many of the brainstormed ideas were difficult to interpret on their own. Presenting the more detailed, annotated design prototypes first was intended to provide judges with a clearer sense of the pupils' design intentions and solution scope, thereby helping them better comprehend the ideation sketches evaluated afterward.

4.2.4 Data analysis

Data were analyzed across three aspects, combining qualitative interpretation with quantitative analysis. An overview of the analysis procedure is presented in Table 4.1. Given the open-ended nature of the design task, we first categorized the designs and documented the frequency of each type. We then examined whether certain types of designs were statistically more likely to be favored by judges. Second, we analyzed judges' comments made during comparative judgement to understand the criteria they used to justify their decisions. The comments were imported into MAXQDA 2022 for qualitative coding. An inductive, data-driven approach was used to generate an initial coding scheme. A word combination frequency analysis (minimum 2 to maximum 5 words) was conducted to identify frequently referenced phrases, which informed the initial categorization of evaluation criteria. Each sentence in a piece of comment served as a unit of analysis and could be assigned multiple codes.

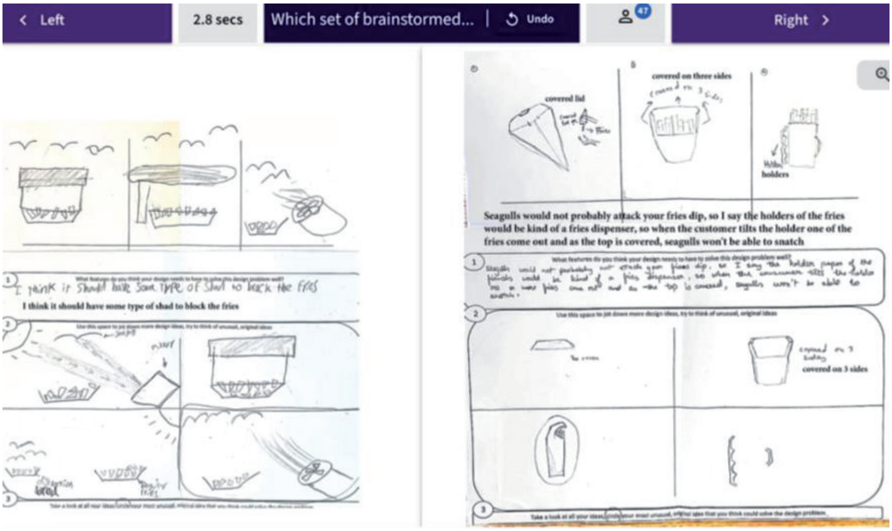


Figure 4.2 A pair of design ideation sketches displayed side by side on the comparative judgement interface

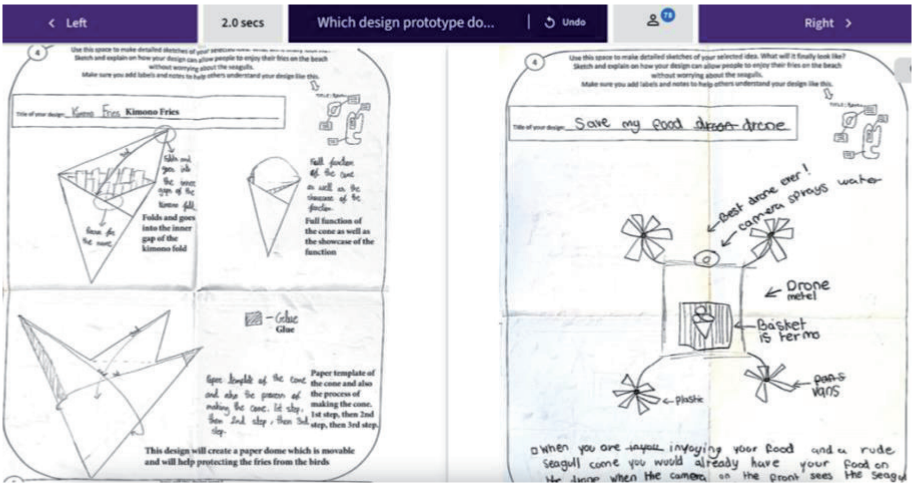


Figure 4.3 A pair of design prototypes displayed side by side on the comparative judgement interface

The coding system was refined iteratively: codes were renamed, merged, or split into sub-codes to more accurately reflect the range of criteria and rationales expressed in the comments, with careful consideration of the context in which the comments were made. For example, the code 'user experience' was split into two sub-codes, 'judges' consideration of user experience' and 'pupils' consideration of user experience.' This distinction was intended to separate what judges inferred about user experience based on pupils' designs from pupils' own attention to user-related affordances reflected in their designs. Once the initial coding system was developed, a second coder reviewed and refined it using a set of randomly selected comments in which all initial codes appeared at least once. A third coder then independently coded 5% of the dataset, which was randomly selected using a random number generator. This step aimed to assess inter-coder agreement and further refine the codebook. Any discrepancies were discussed until full consensus was reached on code interpretations. The full dataset was then recoded using the finalized codebook.

Third, we examined how judges' use of evaluation criteria evolved as they proceeded through a large number of comparative judgements. Specifically, we analyzed whether the types of criteria applied shifted from early to later stages of the comparative judgement process. To explore this, we visualized changes in judges' use of criteria across different categories by comparing the first and last 15 comments made by each judge. We then tested for statistical differences in their tendency to apply different types of criteria over time.

Table 4.1 Overview of the data analysis procedure

	Analysis 1	**Analysis 2**	**Analysis 3**
Subject of analysis	Pupils' design prototypes	Judges' comments generated during comparative judgement	Judges' use of evaluation criteria over time
Outcomes	Frequencies of design types and statistical comparison of design rankings by types	Inductive coding of comments and iterative refinement of the codebook	Visualization and statistical comparison of criteria used in judges' first versus last 15 comments

4.3 Results

4.3.1 Comparative judgement outcomes and reliability

The average median time spent by the 20 judges for making a judgement was 68.6 seconds for judging design ideation and 73.0 seconds for judging design prototypes. The scale separation reliability (SSR), which is a Cronbach's alpha equivalent in comparative judgement, was 0.834 and 0.753 for judging design ideation and judging design prototypes respectively, both indicating good internal consistency. No critical misfit was found among judges, meaning that each judge's decisions consistently fit with the overall consensus (Pollitt, 2012).

The ideation and prototype scores were generated by uploading the judgement decisions to the Adaptive Comparative Judgement App (Buckley, 2024a, 2024b), where the parameter values were determined by fitting the Bradley-Terry-Luce model using the Supplementary Item Response Theory Models (sirt) package (Robitzsch, 2021) in R. The parameter values indicate the relative ranking of each pupil's design ideation and design prototype works (see Figures 4.4 and 4.5), and will be referred to as design ideation and design prototype scores. Spearman correlation between the parameter values pupils received for design ideation and design prototype was moderate, $r(199) = 0.371$, $p < .01$, 95% CI [0.244, 0.486].

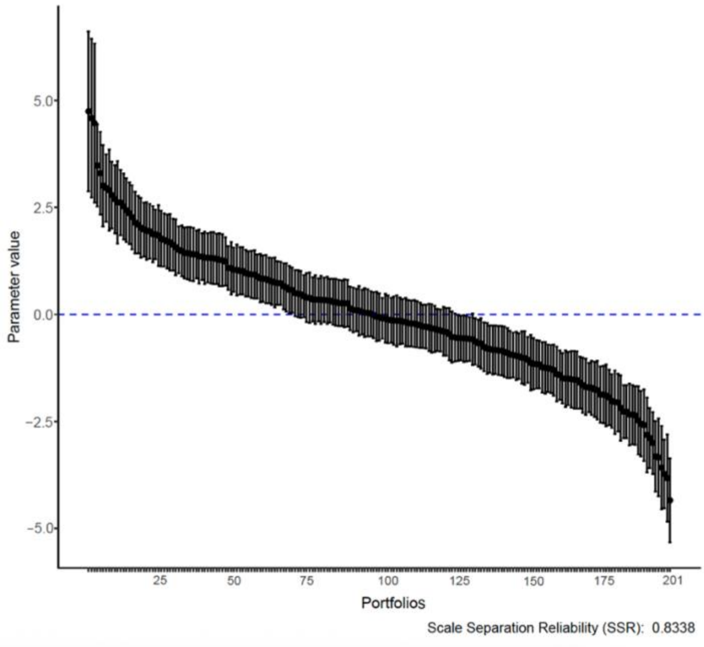


Figure 4.4 Comparative judgement ranks for pupils' design ideation

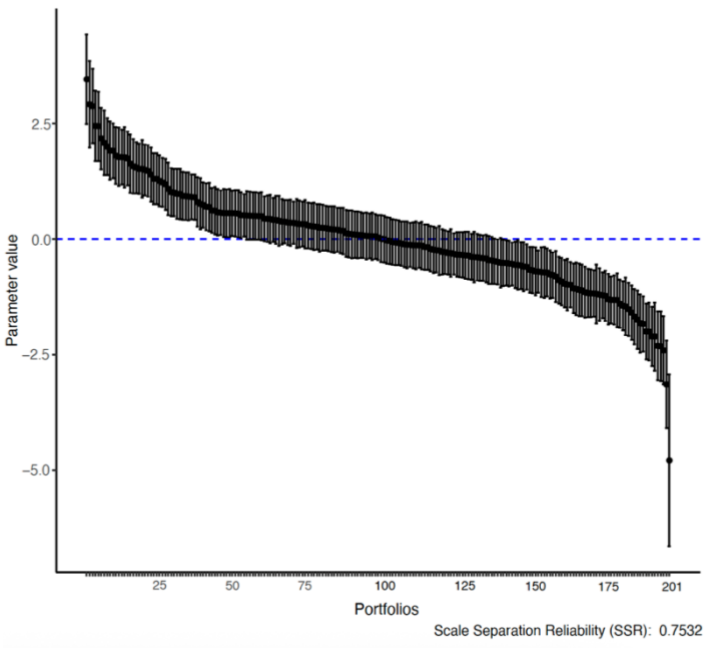
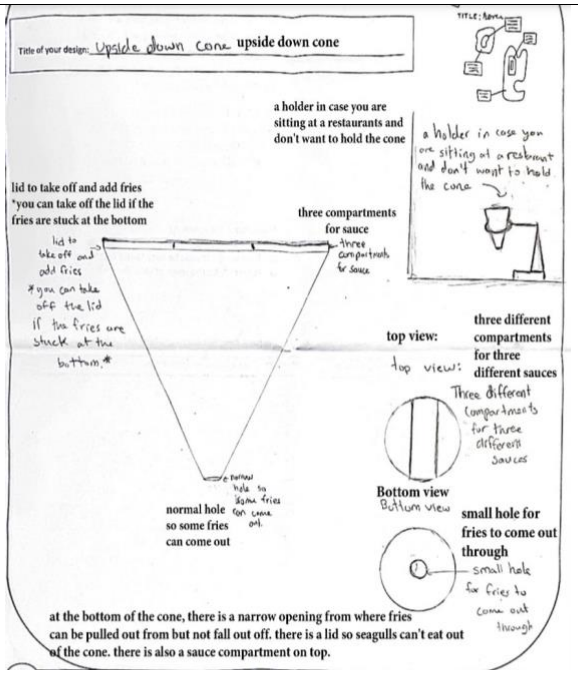


Figure 4.5 Comparative judgement ranks for pupils' design prototypes

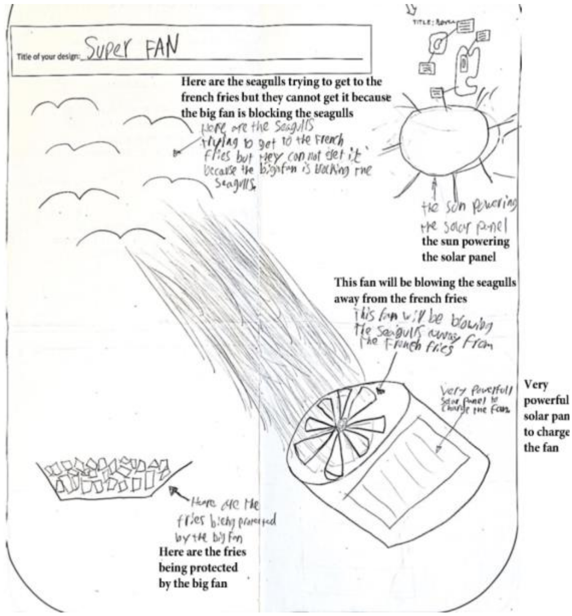
4.3.2 Types of designs developed

We observed that several judges attempted to categorize the design ideas during evaluation—either to facilitate their comparative assessment process or because comparing designs across different conceptual categories made decision-making more challenging. To illustrate the breadth and diversity of pupils' designs, we categorized them and further examined whether design scores varied across types. Pupils' design prototypes were classified into six categories: packaging, repellent, shielding, disguise, multi-element system, and distraction. Table 4.2 presents the percentage of designs in each category, along with representative examples. Designs that clearly demonstrated one or two key features were categorized based on their most salient characteristics. Designs that integrated three or more distinct features spanning multiple categories were classified as multi-element system designs.

Table 4.2 Overview of types of design prototypes developed by pupils (n = 201)

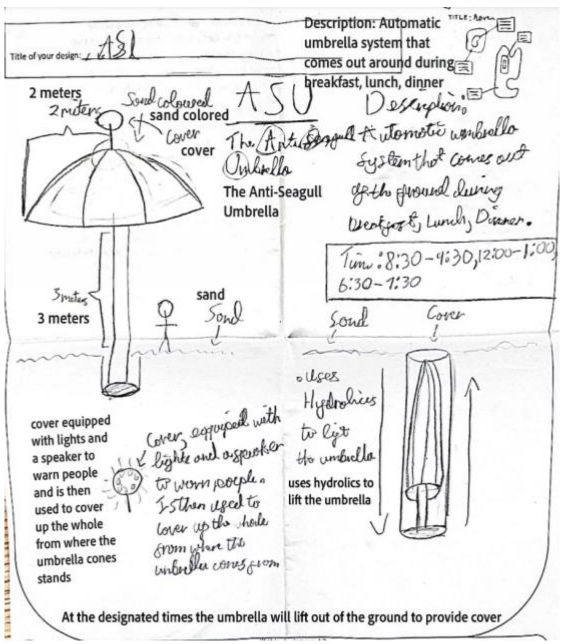
Types of design	**Percentage**	**Design examples with typed annotations**
Packaging	41.29 %	 An upside-down fries packaging that is covered on the top and has users pull fries from the narrow opening at its bottom

Repellent (physical/ chemical) 24.38 %



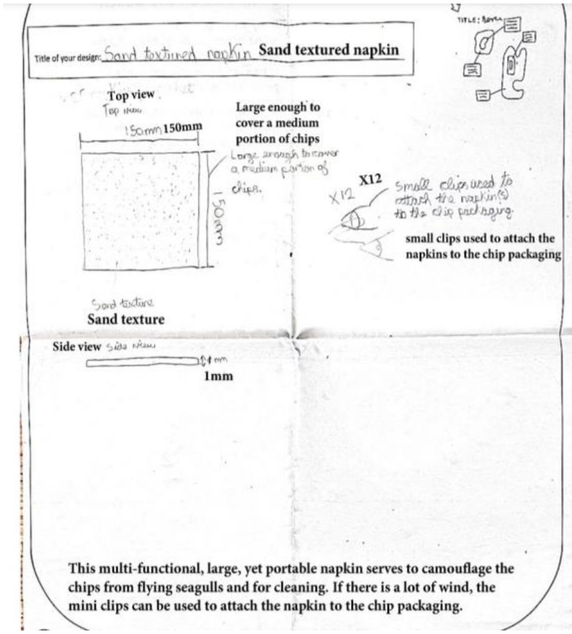
A solar-powered fan to blow seagulls away from the fires

Shielding 13.43%



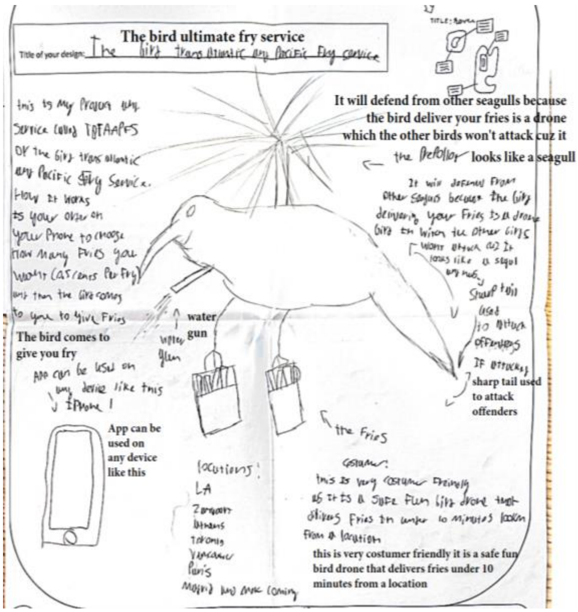
An automatic umbrella system that is lifted out of the ground during mealtimes to shield people from seagulls

Disguise 7.96%



An attachable, sand-textured napkin that camouflages fries from seagulls and can be used for cleaning after eating

Multi-element system 7.46%



A fries delivery service that delivers fries to users by using seagull-shaped drones to deter real seagulls

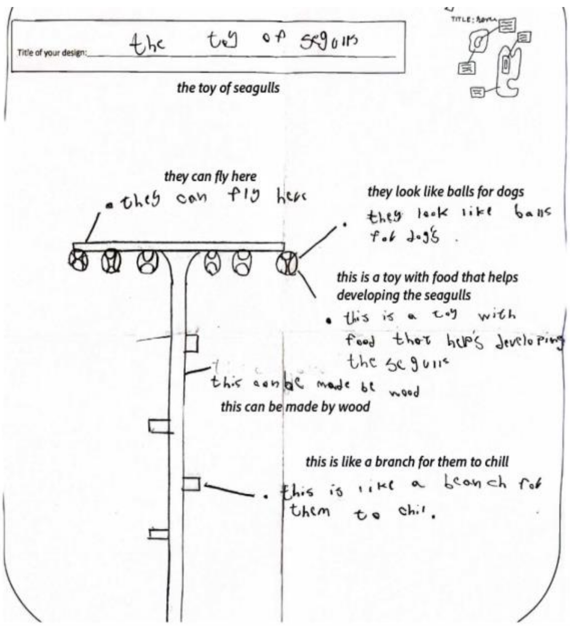
Distraction	5.47%	 A toy for seagulls with food inside that attracts seagulls there

Table 4.2 continued

Based on the six types, we conducted another comparative judgement analysis by computing judges' decision data based on the types of design that were being compared. For example, instead of having 'prototype 1' versus 'prototype 2' in the decision data, where 'prototype 2' was chosen as the winner, we reformatted the data as 'Packaging' versus 'Distraction', where 'Distraction' was chosen as the winner. Results are presented in Figure 4.6.

There appeared to be differences in how certain types of designs were ranked, with some types receiving higher ranks than others. To examine whether such differences were statistically significant, a Kruskal-Wallis H test was conducted in SPSS. A significant difference was found across design types, $H(5) = 13.67$, $p < .05$, providing evidence that certain types of designs were considered to be more or less creative than others. Follow-up post hoc pairwise comparisons (unadjusted p-values) revealed that packaging designs were rated significantly lower than both multi-element system designs and distraction designs ($p < .05$) and that repellent designs were also rated significantly lower than multi-element designs and distraction designs ($p < .05$). To account for potential inflation of Type I error due to multiple comparisons, significance levels were adjusted using Holm-Bonferroni correction (Holm, 1979) with an alpha level of .05. After adjustment, none of the pairwise differences remained statistically significant. These results suggest that while overall group differences exist, the differences may be subtle and distributed across design types rather than driven by a single pairwise comparison difference. It is also possible that the effect sizes of these differences were relatively small (Buckley, 2024c), resulting in limited statistical power to detect differences at the pairwise level.

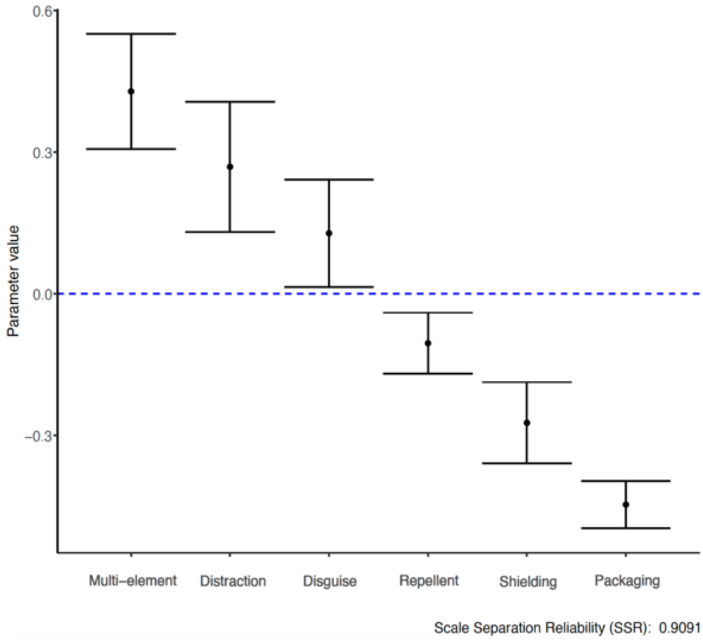


Figure 4.6 Comparative judgement ranks by types of design

4.3.3 Qualitative coding of judgement criteria

Each judge provided at least 30 comments separately for evaluating design ideation and design prototypes, resulting in a total of 600 pieces of comments for each design phase. While all judges were instructed to leave comments for the first and last 15 judgements they made, some judges provided additional comments throughout the process. To ensure that we analyzed a comparable number of comments across judges, we included only the first and last 15 comments provided by each judge in this analysis. This approach was intended to capture judges' insights from both the beginning and the end of the judging process.

On average, each comment on design ideation was 34 words in length, and each comment on design prototypes averaged 44 words. Two coders achieved good inter-rater agreement when assigning codes to selected comments regarding design ideation (77.9%) and design prototypes (79.9%). Any discrepancies were discussed until full agreement on code interpretations was reached. Through iterative coding, a total of 39 codes were developed to capture judges' evaluative criteria. These codes were grouped into six overarching categories: Novelty, Idea qualities, Usability, Feasibility, Presentation, and Problem-solving. An additional category, Idea generation, was developed to capture aspects specific to evaluating design ideation. Table 4.3 presents the definitions for each category and lists the codes included under Novelty, which was of particular interest in this study. A complete description of the remaining codes can be found in the codebook provided in the Appendices.

Table 4.3 Qualitative coding of judges' evaluative considerations: categories and codes

Code categories	Definition	
Novelty	This category refers to how uncommon, unique, or unconventional the design (or idea) is.	
Usability	This category refers to how well the design (or idea) supports users' needs and ease of interaction.	
Presentation	This category refers to how well the design (or idea) is communicated and visually represented.	
Idea qualities	This category involves a series of subjective attributes of the design (or idea).	
Feasibility	This category refers to the practicality of bringing the design (or idea) to life.	
Problem-solving	This category refers to how comprehensively the design (idea) addresses the design brief.	
Idea generation	This category refers to the variety and quantity of ideas produced during the ideation phase.	
Novelty codes	**Definition**	**Examples**
Original/unique	Judges commenting on a design being innovative, new, original, unique, rare, fresh, not often seen, or proposing a new concept; also on a design being an obvious idea, similar to other frequently seen or existing solutions	"It stands out as an original idea compared to the more straightforward container solution" "Ideas on the left are a bit boring because it is obvious"
Uncommon mechanism (mechanical/ structural)	Judges commenting on an uncommon mechanism—a specific structure or mechanical component—in the design, including the shape, the structure, the compartments and components, the way of assembling/attaching parts of the design, or the motion or interaction generated due to these compartments, that are unusual, novel, or not seen before	"The idea of linking all these elements and make them activate in sequence is really unique" "left design is more innovative using rubber mouth-like openings so it can keep original state automatically"
Creatively combining different ideas	Judges commenting on a design being a creative and meaningful combination of different concepts that may have otherwise been common ideas	"I like that the left one uses play predators in combination with the sound of playing kids" "I think the left design is a bit more out-of-the-box thinking with the use of water and sound to repel the seagulls."

Different starting point	Judges commenting on a design taking a different direction or perspective, changing the context, or reframing the scope of the problem; could be expressed in these forms: "instead of (a common way) the child did (new and different) way" "it's not about a (common idea)...but about...."	"the child did not think about the fries or the packaging but rather what the seagulls really want which in the end is just food" "The child focuses on the delivery of the fries and not the packaging"
Modifying an otherwise usual idea	Judges commenting on the core design idea being common, but noting that the creative features added to the design made it distinct from the otherwise conventional ideas or types of usage	"they also thought of the bottle spraying based on a timer, which would make it a bit more innovative than already existing scent sprays" "While the core lies in covering the fries the creativity lies in the shape and the deception through a baby crib"
Creative/ out-of-the-box (general)	Judges commenting on a design being 'creative' or 'out-of-the-box' in a general way that does not fall into the above-mentioned categories, that is, without mentioning that it is new and rare, exhibiting uncommon mechanisms, combining different ideas, taking a different direction, or modifying a usual idea	"in the end I think the left one is a bit more outside of the box thinking" "the drawing of the seagull is a creative idea"

Table 4.3 continued

The frequencies of the main code categories are shown in Figures 4.7 and 4.8. Codes under each category are listed below in Table 4.4, along with their frequencies relative to the total number of code appearances ($n_{\text{Ideation}} = 1720$, $n_{\text{Prototype}} = 1935$).

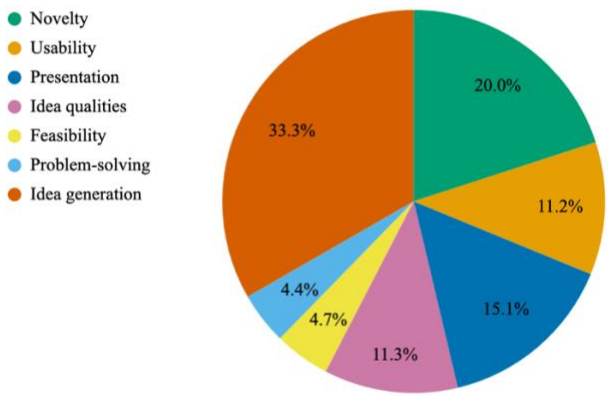


Figure 4.7 Percentage of main code categories identified in design ideation comments

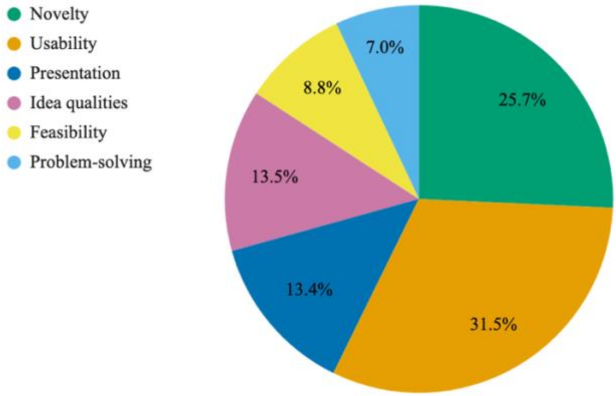


Figure 4.8 Percentage of main code categories identified in design prototype comments

Table 4.4 Qualitative coding of judges' evaluative criteria: categories and codes

Category	Codes	Code frequency in design ideation comments (total code appearances = 1720)	Code frequency in design prototype comments (total code appearances = 1935)
Novelty	Original/unique	0.099	0.094
	Creative/out-of-the-box (general)	0.062	0.057
	Uncommon mechanism	0.018	0.027
	Creatively combining different ideas	0.013	0.027
	Different starting point	0.008	0.037
	Modifying an otherwise usual idea	—	0.015
Usability	Useful & functional	0.039	0.102
	Simple & intuitive	0.016	0.045
	Pupils' consideration of user experience	0.024	0.045
	Judges' consideration of user experience	0.010	0.039
	Multiple elements/features	0.013	0.038
	Multiple functions/purposes	0.006	0.022
	Tailored for the context	0.004	0.013
	Customization	—	0.009
Presentation	Clarity in explanation	0.062	0.075
	Elaborated details	0.080	0.044
	Quality of drawing	0.008	0.014
	Storytelling	0.002	—

Idea qualities	Underdeveloped ideas	0.035	0.013
	Interesting	0.026	0.017
	Fun & playful	0.011	0.026
	Good (general)	0.008	0.019
	Smart	0.008	0.018
	Aesthetics & desirability	0.008	0.016
	Sustainability	0.004	0.012
	Idea potential	0.008	0.005
	Imagination	0.003	—
	Considerate	—	0.005
	Surprising	—	0.003
	Explorative	—	0.002
Feasibility	Realistic to make	0.017	0.035
	Involving technology	0.018	0.020
	Considering materials	0.009	0.019
	Cost-effectiveness	0.003	0.010
Problem-solving	Thought-through solutions	0.026	0.045
	Meeting the design brief	0.019	0.025
Idea generation	Diverse directions	0.243	—
	Quantities of ideas	0.052	—
	Variations of a key idea	0.038	—

Table 4.4 continued

4.3.4 Shifts in criteria use from the beginning to the end of the judging process

The third analysis examined whether, and how, the frequency of judges' use of different criteria changed between the start and the end of the judging process. For each judge, we calculated the relative frequencies of criteria used within each code category based on their first 15 and last 15 recorded comments. Figures 4.9 and 4.10 visualize the shifts in judges' use of different evaluative criteria over time.

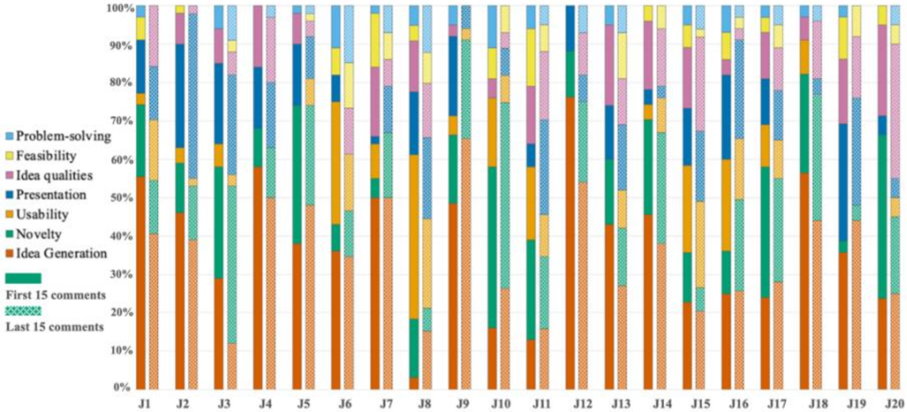


Figure 4.9 Relative frequencies of criteria from seven code categories in the first and last 15 design ideation comments from 20 judges

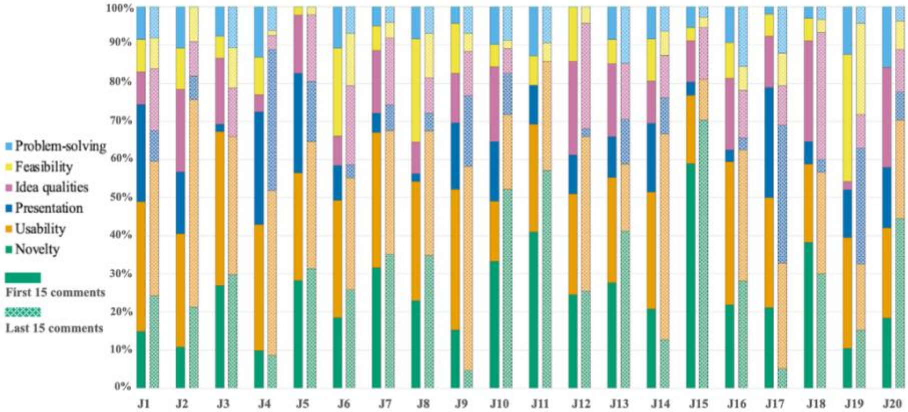


Figure 4.10 Relative frequencies of criteria from six code categories in the first and last 15 design prototype comments from 20 judges

To examine whether the 20 judges' use of evaluation criteria changed significantly between their first and last 15 comments left during comparative judgement, we conducted non-parametric Wilcoxon Signed-Rank Tests in SPSS. No significant differences were found in the frequency of applying criteria from each category between their first and last 15 comments regarding design ideation. For design prototypes, judges appeared to use Novelty criteria (Z = -2.092, p = .036, $p_{adjusted}$= .180) more frequently and Feasibility criteria less frequently (Z = -2.620, p = .009, $p_{adjusted}$= .054) in their last 15 comments. However, the adjusted p-values (Holm-Bonferroni correction) suggested that these differences were no longer significant. Overall, the judges applied the coded criteria with largely comparable frequency at both the beginning and the end of the judging processes for both design ideation and design prototypes.

4.4 Discussion of main findings

This study illustrated how 201 pupils aged 10 to 14 generated a range of creative designs—

packaging, shielding, physical or chemical repellent, disguise, distraction, and multi-element systems—to address a given design task. We explored the decisions and evaluative considerations provided by 20 industrial design students acting as judges, who used comparative judgement to assess both pupils' design ideation outcomes and design prototypes. The main findings, outlined below, are discussed by merging qualitative and quantitative results (Fetters et al., 2013).

4.4.1 Judges' decisions and types of designs

Our analysis of the design types produced by pupils revealed overall differences in how these types were ranked by the judges. This finding is not unexpected, as several judges noted that certain types of design were more creative than others. For example, Judge 3 remarked, "I like that it is not a packaging design but something to distract the seagulls and keep them away due to the movement," and Judge 7 mentioned, "the designer thought out-of-the-box when designing a hat instead of a fry packaging which is good!" Packaging design was the most frequently attempted solution (41.29%), whereas other types, such as distraction designs, were relatively rare (5.47%). Since judges were explicitly asked to focus on creativity—which implies the need for rare and uncommon solutions—their evaluative considerations appeared to align with their comparative judgement decisions. Statistical analysis using Holm-Bonferroni–adjusted p-values revealed no significant differences in creativity rankings between design types. This suggests that although certain types of design (e.g., packaging) were more prevalent, they were not necessarily viewed as less creative overall by the judges.

4.4.2 Coding scheme and its alignment with existing frameworks

The qualitative coding scheme developed inductively from judges' comments on pupils' designs aligns closely with established frameworks for evaluating product creativity. Specifically, it reflects all three dimensions outlined in the Creative Product Semantic Scale (O'Quin & Besemer, 1989): novelty, resolution (usefulness and effectiveness in solving the design problem), and elaboration and synthesis. It also corresponds with the key criteria identified by Casakin and Kreitler (2005) for evaluating creative design solutions, including innovation, usefulness, and functionality, fulfilling specified design requirements, elaborations, consideration of context, aesthetics, and the number of relevant solutions. Furthermore, our codes align with several core components of the Creative Solution Diagnosis Scale (Cropley & Kaufman, 2012), such as relevance ('solution fits within task constraints'), effectiveness ('the solution does what it is supposed to do'), multiple aspects of novelty ('the solution indicates a radically new approach'; 'the solution offers a fundamentally new perspective on possible solutions'; 'the solution makes use of new mixtures of existing elements'; 'the solution helps the beholder see new and different ways of using the solution') (p. 124), as well as how complete, pleasing, and sustainable the solution is. Lastly, the additional coding category developed for judges' comments on design ideation aligned with the key criteria for effective design ideation—variety and quantity of ideas—identified by Shah and colleagues (2003). Collectively, these alignments support previous findings that leveraging expert judgement can enrich the conceptualization of the assessed construct (Leisterhuis et al., 2022; Whitehouse, 2012).

Compared to previous studies that reported creativity or innovation as low-frequency criteria in the comparative judgement of design quality, our study took a different approach. Through explicitly instructing pupils to make creative designs and asking judges to prioritize creativity in their evaluations, Novelty was positioned as a central consideration in the assessment process. Nevertheless, considerations beyond Novelty—such as Usability, Presentation, and Idea Qualities—still played an important role in judges' decisions. In the evaluation of design ideation, judges' comments were most frequently coded under Novelty and Idea Generation, which together accounted for more than half of all codes. In contrast, Usability and Feasibility were referenced less frequently, aligning with the primary aim of design ideation—to encourage thinking outside the box without placing too much emphasis on practicality.

In the evaluation of design prototypes, judges referred more often to Usability ($f = 0.315$) than to Novelty ($f = 0.257$). This aligns with the well-established view that creative products need to be both novel and effective (Cropley & Kaufman, 2012; Horn & Salvendy, 2006; Runco & Charles, 1993; Sarkar & Chakrabarti, 2011). Technological design, in particular, requires relevant and workable solutions, as novelty alone is insufficient (Cropley & Cropley, 2010). Interestingly, our coding of judges' comments revealed that clearly distinguishing between what is considered 'creative' and what is 'useful' was not always possible—or even necessary. Several judges highlighted that a design's affordances contributed to its perceived creativity. For example, Judge 13 remarked, "It seems multifunctional and has utilities which consider the environment like being able to stand in the sand and the context-aware solution is interesting and creative." Similarly, Judge 18 noted, "The right one is a bit more creative to me because it does more than blocking the fries from the seagulls it scares them away as well." These comments illustrate how judges engaged in holistic comparisons between solutions, blending multiple criteria in their assessment of creativity. This underscores a key advantage of comparative judgement over rubric-based scoring: judges are not burdened with the need to precisely differentiate between the desired qualities (Bejar, 2012; Pollitt, 2012) or assign discrete scores to separate criteria. Instead, they can draw on multiple considerations in a natural and intuitive way.

In summary, whereas conventional scoring methods rely on predefined criteria and emphasize fine-grained distinctions between criteria—even when the differences are subtle—comparative judgement allows evaluators to integrate various considerations into a cohesive judgement. Furthermore, we propose that comparative judgement may be particularly well suited to capturing insights that emerge from judges weighing multiple evaluation criteria—insights that, as Sadler (2008) noted, traditional assessment rubrics might overlook.

4.4.3 Comparison of the first and last 15 comments

As in design problem-solving, where tentative solutions and problem context co-evolve through ongoing reinterpretation and redefinition, resulting in a dynamic understanding of the design context (Dorst & Cross, 2001), judges' conceptualization of design quality can also shift as they make successive comparisons. Rubric-based assessment is susceptible to rater drift, where raters deviate from their original scoring standards over time, leading to inconsistencies and reduced reliability across the assessment process (Hoskens & Wilson, 2001; Myford & Wolfe, 2003). It is therefore important to examine whether judges' evaluative criteria in comparative judgement may likewise change over the course of the evaluation process.

Due to the random pairing of pupils' work, the designs that judges encountered at the beginning and end of the judging process likely differed, which may explain any observed changes in their rationale. However, with 20 judges each providing 30 pieces of comments—detailing their rationales for nearly one-third of their judgements—the sample size was sufficiently large to mitigate these variations and allowed us to analyze shifts in the criteria used, an aspect rarely reported in previous comparative judgement studies.

After Holm-Bonferroni correction, none of the Wilcoxon tests comparing the criteria judges used in the first and last 15 comments reached significance for either the ideation or prototype phase. These results indicate that judges' reliance on the various evaluation rationales remained largely stable from the beginning to the end of the assessment. Nevertheless, non-significant results do not confirm the absence of change. Both the visualizations and statistical analyses indicated that, by the end of their evaluation of design prototypes, judges referred more frequently to Novelty and less to Feasibility compared to the beginning. This may reflect a growing focus on the primary evaluation goal—creativity—as judges became more acquainted with the comparative judgement process. For instance, Judge 13 commented, "Irrespective of which is the better solution overall (it might be Left) the Right one feels more creative." Holistically evaluating the design means constantly considering trade-offs between criteria, such as between novelty and feasibility. On one hand, comparative judgement offers an integrated perspective that traditional methods of grading may not fully reflect. On the other hand, synthesizing multiple, and sometimes competing, evaluative considerations may be cognitively demanding for some judges (Forthmann et al., 2017).

It is also worth noting that the relative frequency of judges using different criteria (as shown in Figures 4.9 and 4.10) does not necessarily indicate that they made different decisions. Judges could favor the same design but justify their choices using different rationales. For example, Judge 1 praised a design's creativity due to its details and that "all steps and the goal of the design are explained." Meanwhile, Judge 20 commented on its functionality and novelty—"Although there is a scope that this could get attacked by other seagulls to get the fries while it is delivering, the idea looks very distinct." Despite the differing justifications, both judges—and many others—agreed on the overall quality of this design. This highlights how, in comparative judgement, judges can provide different but valid rationales while still achieving good inter-rater reliability.

4.4.4 Reliability among judges and challenges in decision-making

The scale separation reliability (SSR)—which essentially measures how all judges collectively produce a similar ranking of design works—was found to be good in this study and comparable to previous research (Bramley & Vitello, 2019; Jones & Alcock, 2014; Verhavert et al., 2019). It is important to note that the pairing of pupils' work in our study was random rather than adaptive, whereas studies that reported even higher SSR of above 0.90 were predominantly using algorithms that adaptively present pairs of works with a similar number of wins for more efficient comparison (e.g., Bartholomew et al., 2020; Buckley et al., 2022).

The qualitative coding of judges' comments revealed several challenges during the evaluation process. Some expressed difficulty when the paired works were similar, as Judge 19 noted, "It is a bit hard to judge and compare the two concepts because the ideas are not very different; the concept of working on the packaging in different ways has more creativity." Conversely, when the paired works were markedly different, judges faced the challenge of balancing various attributes. As Judge 16 explained, "left has more practical creativity while right has more imaginative creativity… hard to judge…will go with left since some features like the 'dark color to keep heat' help enhance the experience." Future studies may want to explore how judges deal with potentially conflicting rationales, and whether targeted training, evaluation guidelines, or improved pairing algorithms could help reduce the cognitive load during judgement.

Another potential factor influencing SSR was the shared background of the judges. All judges were trained in the same industrial design department, which may have contributed to a high degree of consensus, particularly when evaluating design ideation. A frequently cited criterion was the quantity of ideas. For example, Judge 1 noted, “Right shows more ideas. Quantity will lead to quality during brainstorming, therefore it is more creative.” Similarly, Judge 4 commented, “As you need a lot of ideas in order to come up with the real creative ideas, I see more potential for people who are able to generate a lot of ideas.” While research often supports the claim that quantity breeds quality (Kudrowitz & Wallace, 2013; Paulus et al., 2011), other studies have challenged this notion. For example, Goldschmidt and Tatsa (2005) found no correlation between design quality and the sheer number of ideas generated, as many ideas do not meaningfully address the design problem and would likely be discarded after the ideation phase. Therefore, while consensus among judges could enhance consistency in decision-making, it also raises concerns about potential bias stemming from homogeneity in evaluative perspectives.

4.5 Limitations and future directions

One limitation of using comparative judgement as an assessment tool lies in its focus on relative, rather than absolute, quality among the compared works. Since judges were asked to select the more creative design from each pair, it is possible—as some judges noted in their comments—that neither design was particularly creative. Accordingly, it is important to clarify that our outcomes reflect the relative creativity observed within the studied sample of pupils’ design work. While this is consistent with the foundational principle of comparative judgement, it may be viewed as a limitation, as the relativity of the judgements could restrict the generalizability of our results—particularly in comparison to results derived from standardized rubrics. However, this limitation pertains more to our study design than to the comparative judgement method itself. Recent research has proposed strategies to transform the produced relative rank order into grades, which could then be compared across assessments (cf. Egelandsdal et al., 2025). Moreover, methods have been developed to merge and compare independent comparative judgement rank orders (Benton, 2021; Buckley et al., 2023; Buckley & Canty, 2022; Verhavert et al., 2022), further extending the applicability of this method in broader assessment contexts.

A second limitation concerns the composition of the judging panel. All judges involved in this study were master’s-level design students from a single university department, which may have shaped a shared conceptualization of creativity. Previous research suggested that judges’ view of design quality can be influenced by their cultural backgrounds (Bartholomew et al., 2020) and levels of professional experience (Strimel et al., 2021). As this study focused on evaluating the creative design work of primary and secondary school pupils, the absence of school D&T teachers on the judging panel limited the educational relevance of our findings. Future research could address this by involving a more diverse panel of judges—including professional design practitioners and D&T educators—to gain a more comprehensive understanding of creativity in pupils’ design. In addition, emerging research highlights the potential of involving learners themselves in comparative judgement as a means of providing formative peer feedback and enhancing learning (Bartholomew et al., 2019; Bartholomew et al., 2022). Building on this, future studies could explore whether engaging pupils in comparative judgement helps foster their understanding and development of creativity in design.

4.6 Conclusion

By analyzing a total of 1200 comments provided by judges, this study unpacked the key rationales guiding their comparative judgement of pupils' creative design work. Our results revealed a consistent use of evaluative criteria across the judging process, with judges applying similar criteria from their initial to final comparisons. These criteria, identified through qualitative coding, aligned closely with established frameworks for assessing product creativity. Our findings support the growing body of research attesting to the validity and reliability of comparative judgement as an assessment method. Beyond this, comparative judgement appeared especially useful in capturing nuanced insights when assessing complex constructs such as creativity and design quality. Judges naturally integrated multiple evaluative criteria to form holistic decisions. These findings suggest that comparative judgement is not only a feasible alternative to rubric-based assessment but also offers rich insights for evaluating creative, open-ended tasks. Future research could further examine how judges balance trade-offs among criteria, and how optimized pairing algorithms might help reduce cognitive load while maintaining judgement quality. Additionally, involving a broader range of evaluators—including educators and learners themselves—may yield a more comprehensive understanding of creativity in the context of design education.

CHAPTER 5
Spatial ability’s contribution to pupils’ creative design performance: Differential roles in ideation and prototyping

CHAPTER 5
Spatial ability's contribution to pupils' creative design performance: Differential roles in ideation and prototyping

Prior research examining the relationship between spatial ability and creative design performance among college students has yielded mixed findings. Even less is known about how spatial ability predicts creative design performance in younger pupils and how its role compares to other cognitive abilities commonly linked to design performance, such as creativity. This study examines how spatial ability, divergent thinking, and creative imagery abilities predict design performance across two stages—ideation and prototyping—in a sample of 201 pupils aged 10 to 14. Results revealed that spatial ability correlated significantly with design ideation, design prototype, and divergent thinking scores, but not with creative imagery. Using hierarchical regression analyses, we examined the unique contributions of each cognitive skill to design ideation and design prototype performance. Creative imagery abilities significantly predicted design ideation scores after controlling for age and verbal ability, whereas divergent thinking and spatial ability did not. However, in predicting design prototype performance, spatial ability emerged as a significant predictor in addition to creative imagery abilities and divergent thinking. Mediation analysis further revealed that creativity did not meaningfully mediate the relationship between spatial ability and either design ideation or prototype scores. Taken together, our results suggest that while creative imagery abilities are important for generating creative design ideas during ideation, spatial ability plays a distinct role over and above creativity as pupils develop initial ideas into sketched prototypes.

5.1 Introduction

Design—both as a method to purposefully explore and solve problems, and as a practice to better our surroundings—is integral to the development of society and the improvement of daily life (Custer, 1995; Gero, 1990). It is a dynamic, iterative process that transforms ill-defined, ambiguous problems into solutions through the creation and refinement of conceptual, visual, and tangible representations (Lawson, 2005). In K-12 education, design appears both as a distinct subject central to the Design & Technology discipline (Kimbell & Stables, 2007) and as an integrative link among science, technology, engineering, and mathematics (STEM) problem-solving (de Vries, 2021; Hallström & Ankiewicz, 2023). Design thinking and practices have received growing recognition in STEM education worldwide (Goktepe Yildiz & Ozdemir, 2020; Ng & Chan, 2019; Ramey & Uttal, 2017; Simeon et al., 2022) for their engaging and versatile nature, which enables the integration of real-world problems such as complex social and environmental issues into the classroom (Kimbell & Stables, 2007). To advance effective design education and practice, it is crucial to develop a better understanding of the cognitive processes underlying successful design performance.

At its core, design is a creative activity (Christiaans, 2002; Cropley & Cropley, 2010; Dorst & Cross, 2001; Goldschmidt, 2016; Lawson, 2005). The role of creativity is evident across disciplines, including engineering design, industrial design, and architectural design (Casakin & Wodehouse, 2021) and across all phases of the design process (Howard et al., 2008; Tassoul, 2009). Developing a creative design involves multiple iterative steps, such as defining and framing the design problem (Dorst & Cross, 2001), generating ideas divergently and evaluating ideas convergently (Goldschmidt, 2016), and prototyping and iterating on selected concepts (Rosenman & Gero, 2013). In these processes, a variety of cognitive skills are at play, such as mental simulation (Christensen & Schunn, 2009) and imagery (Athavankar, 1997; Tedjosaputro et al., 2018), eye-mind-hand coordination skills (D'souza, 2010), and spatial skills (Shah et al., 2013). To think spatially is to create, transform, and update mental and physical representations of not only concrete objects but also abstract ideas (Tversky, 2015). In this sense, abstraction—which gives rise to ambiguity, an element that Tversky viewed as a key driver of creativity in design—is inherently spatial.

Spatial ability has played a vital role since early humans began developing artifacts and tools, and it continues to drive scientific and technological innovation today (Ferguson, 1977; Kell et al., 2013). Given the well-established link between spatial ability and STEM learning (Buckley et al., 2018; Wai et al., 2009; Uttal et al., 2013), and the importance of cultivating spatial ability among K-12 pupils (Newcombe, 2013; Uttal & Cohen, 2012), there is a growing need to understand how spatial ability contributes to performance in other disciplines—particularly design, which has received comparatively less attention in spatial ability research. Although prior research has suggested evidence linking creativity measures to design performance (Casakin et al., 2010; Charyton et al., 2011; Kokotovich & Purcell, 2000), the role of spatial ability in creative design performance remains less well understood. Empirical findings have been mixed: some studies found a positive correlation between spatial ability and creative design performance (Chang, 2014; Cho & Suh, 2019; Suh & Cho, 2020), whereas others reported no significant relationship (Allen, 2010; Cho, 2017; Reid & Sorby, 2023). In parallel, a growing body of intervention research demonstrated gains in spatial skills through design-related activities, suggesting a possible causal connection between design learning and spatial skills development. However, these studies primarily focused on technical drawing and graphic-based practices (Sorby, 2009; Sutton & Williams, 2010), computer-aided design (Dere & Kalelioglu, 2020; Guo et al., 2022), or engineering design (Goktepe Yildiz & Ozdemir, 2020; Raju et al., 2023). These contexts often emphasize modeling and meeting concrete design constraints rather than developing creative design ideas and prototypes. Taken together, the mixed and sometimes indirect findings—drawn largely from studies with higher education students—highlight the need to clarify the relationship between spatial ability and creative design performance, particularly among under-studied, younger age groups.

The current study addresses these gaps by examining the relationship between pupils' spatial ability and their creative design performance, while also accounting for established measures of creativity. Specifically, we focus on pupils in the final year of primary school and the early years of secondary school, when design is typically introduced as a formal component of the curriculum (International Baccalaureate Organization, n.d.). Analysis of cognitive development trajectories (Claxton et al., 2005; Smith & Carlsson, 1983) and neuroimaging data (Kleibeuker et al., 2013) suggests that late childhood and adolescence are key periods for the development of creative capacities, driven by growing cognitive flexibility. Engaging in design—a practice of creativity grounded in real-world contexts—therefore aligns with the developmental needs and skills of pupils in these age groups. As Coxon (2012) highlighted, creativity and spatial ability are two 'allies in innovation' that deserve greater attention in today's talent development and educational policy agendas. Accordingly, gaining insight into the unique and combined contributions of creativity and spatial ability to design outcomes can help clarify the cognitive processes underpinning creative design.

5.2 Literature Review

5.2.1 Design as a creative and cognitive endeavor

What separates a creative design from a routine design is how it stands out from existing products and expands the solution space (Rosenman & Gero, 2013). Creativity in design can be understood as transforming the familiar into the unique—whether perceptually, physically, conceptually, functionally, or through a combination of these dimensions (Suwa et al., 1998). Creative design is therefore not the result of an "Aha" moment, but rather emerges through a series of deliberate, research-informed processes (Howard et al., 2008; Lawson, 2005). For example, designers employ creative techniques to draw inspiration from distant analogies (Tassoul, 2009), as well as prompts and scaffolds to help overcome design fixation (Chrysikou & Weisberg, 2005). As a result, designers purposefully develop innovative solutions that are not only novel but also effective (Cropley & Cropley, 2010). These structured processes share resemblances with how innovative solutions are developed in fields such as engineering and technology (Custer, 1995; Howard et al, 2008). However, unlike fields that prioritize functionality and feasibility, design prioritizes functional creativity (Cropley & Kaufman, 2012) and retains space for artistic expression and creativity (de la Harpe et al., 2009), making it an ideal form of learning for K-12 pupils to develop creative works that carefully consider both novelty and usefulness (Bartholomew et al., 2020; Rutland & Barlex, 2008).

The emphasis on purposeful creativity in design aligns with broader psychological theories that conceptualize creativity as creative potential (Runco, 2010; Runco & Acar, 2012). This potential is commonly operationalized through divergent thinking, which refers to an individual's ability to produce multiple, original, and varied ideas. Divergent thinking is typically assessed using psychometric tasks such as the Alternate Uses Tasks (AUT; Guilford, 1967) and the Torrance Tests of Creative Thinking (TTCT; Torrance, 1974). Performance on these tasks has been found to positively associate with real-world creative behaviors and achievement (Jauk et al., 2014; Kim, 2008; Runco & Acar, 2012). Although divergent thinking tasks capture an important component of creative ideation, they primarily emphasize idea fluency and originality under time constraints and may not fully reflect the complexity of real-world creative problem solving. Empirical findings on the relationship between divergent thinking and creative design performance have been mixed. For example, Casakin et al. (2010) found that architectural design students' general creative thinking abilities were moderately correlated with their real-life problem-solving performance in architectural design. Similarly, using a combination of quantitative and qualitative analysis, Atkinson (2000) reported an overall positive association between 15-16-year-olds' design and technology project performance and their divergent thinking. In contrast, Cho (2017) found no significant correlation between first-year students' TTCT scores and their performance in a design studio. One possible explanation for these inconsistencies is that real-world design creativity requires not only the generation of multiple novel ideas but also the ability to evaluate and refine ideas in order to develop effective solutions (Goldschmidt, 2016; Howard et al., 2008). Moreover, psychometric measures of divergent thinking can be influenced by participants' response speed, which does not necessarily correspond to their level of creativity (Jankowska & Karwowski, 2015). As a result, divergent thinking alone may be insufficient to predict performance in authentic design contexts. These mixed findings indicate the need for a more nuanced understanding of how divergent thinking relates to creative design performance among school-aged pupils.

Creative imagery represents another important dimension of creative cognition. Whereas divergent thinking tasks emphasize verbal or figural idea generation, creative imagery refers more specifically to the ability to visualize, manipulate, and transform vivid mental images in original ways (Finke, 1990; Runco et al., 1998). Creative imagery has played a key role in scientists', engineers', designers', artists', and technicians' making of creative discoveries and inventions throughout history (Ferguson, 1977). Empirical studies showed that university students with design training produced nearly twice as many creative mental synthesis ideas in a creative imagery task as non-design students (Kokotovich & Purcell, 2000), and that experienced designers effectively used mental imagery to envision and iterate on complex product designs even when blindfolded (Athavankar, 1997). Furthermore, creative imagery abilities, rather than divergent thinking, were found to predict five-year-olds' ability to construct mental models of space (Jankowska et al., 2019). Thus, while creative imagery is related to other measures of creative potential (Jankowska & Karwowski, 2015), such as AUT and TTCT, it represents a partially distinct construct that emphasizes the generation and transformation of vivid mental representations (Dziedziewicz & Karwowski, 2015). This conceptual differentiation highlights the value of further investigating whether and how creative imagery predicts young pupils' creative design performance.

5.2.2 Spatial ability and its relevance to design

Much research has positioned design as a cognitive activity (Ball & Christensen, 2019; Gero, 1990; Goldschmidt, 2016; Suwa et al., 1998), drawing attention to the role of cognitive skills in shaping design outcomes. Besides creativity, visual-spatial ability is another cognitive ability considered important for design (Athavankar, 1999; Shah et al., 2013; Suwa et al., 1998). A key part of design is envisioning a not-yet-existing configuration to meet the design task's need (Lawson, 2005), which requires designers to mentally visualize its form and function and externally represent them through artefacts. Both qualitative and mixed-methods studies have highlighted the role of spatial thinking in such visualization processes. Expert designers, for instance, make new discoveries by attending to the visual-spatial features in their design sketches (Suwa et al., 1998). The amount of time university engineering design students spend on design processes, such as idea generation and feasibility analysis, appears to depend not only on their level of expertise but also on their spatial ability, with those who exhibited higher spatial ability spending more time analyzing whether their design would work properly (Raju et al., 2023). Novice designers like secondary school pupils also engage spatial thinking by imagining and representing new spatial configurations through talk and gesture (Ramey & Uttal, 2017), or by abstracting form-function relationships and visualizing possible spatial transformations to develop creative, design-by-analogy prototypes (Zhu et al., 2024).

Spatial ability in STEM learning and design-based activities

Our sense of space and spatial representations—such as noticing and interpreting spatial relations between two-dimensional (2D) representations and three-dimensional (3D) objects—develops from an early age (Frick & Newcombe, 2015). Much research emphasizes the value of investigating and fostering spatial ability development from childhood, adolescence, and onwards (Newcombe & Shipley, 2015; Lowrie et al., 2019; Shea et al., 2001; Zhu, Leung, et al., 2023). Across various frameworks, three key dimensions of spatial ability—mental rotation, spatial orientation, and spatial visualization—have been identified as important for learning in upper primary school (Lowrie et al., 2019) and secondary school (Ramful et al., 2017). Drawing on the working definition of spatial ability proposed by Ramful and colleagues (2017), mental rotation involves imagining how a two- or three-dimensional object would appear after being rotated around a point or axis; spatial orientation involves positioning oneself in space and mapping spatial relations across varying scales, locations, and perspectives; and spatial visualization involves mentally visualizing a series of multistep transformations applied to spatial properties of an object. Primary and secondary school mathematics learning, for instance, draws on an interplay of these three spatial skills (Harris et al., 2021; Lowrie et al., 2019; Ramful et al., 2017).

Spatial ability is widely recognized as important for learning STEM disciplines and subdisciplines across K-12 and higher education (Buckley et al., 2018; Newcombe, 2013; Shea et al., 2001; Uttal & Cohen, 2012; Wai et al., 2009). A growing body of research has demonstrated how spatial thinking is engaged in STEM-related projects through design-based practices, such as 9-13-year-olds practicing a range of spatial skills key to engineering learning through design and construction activities (Ramey & Uttal, 2017), seventh-graders exercising visual-spatial thinking and geometric reasoning by learning mathematics in a visual art studio (Kus & Cakiroglu, 2022), and 11-12-year-olds developing spatial understandings of number magnitude through embodied making of data physicalizations (Zhu, Klapwijk, et al., 2023). While these qualitative and mixed-methods studies have provided rich insights into how spatial thinking is engaged in design-based activities, quantitative research could complement these findings by examining the strength of the relationship between spatial ability and various aspects of design, using larger, more diverse samples to enhance the generalizability of findings. Currently, most quantitative research has focused on clarifying the link between pupils' spatial ability and their performance in mathematics or science (Hawes & Ansari, 2020; Hodgkiss et al., 2018). Only a limited number of studies have quantitatively examined how spatial ability relates to K-12 pupils' creative design performance (Chang, 2014). This represents a critical gap, as understanding how spatial ability contributes to creative design could inform both design education and STEM talent development, especially since design is becoming an increasingly prominent approach to integrated STEM problem-solving (Cook & Bush, 2018; Hallström & Ankiewicz, 2023).

Quantitative evidence linking spatial ability and design performance

Evidence from intervention studies provides indirect support for a close connection between design learning and spatial skills development. For example, university students who enrolled in design courses that focused on sketching, modeling, and computer-aided design demonstrated significantly greater gains in spatial visualization and spatial relation skills than students without design training (Lin, 2016). Similar effects have been observed in younger pupils: seven weeks of 3D computer-aided design-based workshops significantly improve sixth graders' spatial visualization, mental rotation, and cross-section visualization performance (Dere & Kalelioglu, 2020), and eight weeks of engineering design activities led to a statistically significant increase in eighth graders' spatial visualization and spatial relations scores (Goktepe Yildiz & Ozdemir, 2020). These findings suggest that design-based activities appear to foster spatial skills development across education levels. However, it remains less clear whether individuals with stronger spatial ability produce higher-quality design outcomes.

Only a limited number of studies have directly examined this relationship in K-12 contexts. For example, Chang (2014) found moderately strong correlations between spatial ability and creative design performance among 12th-grade students completing a computer-aided chair design task. All subscales of spatial ability—spatial orientation, spatial relationship, and spatial perception—correlated positively, with small to medium effect sizes, with novelty, aesthetics, and functionality aspects of students' creative design.

Among university students, findings have been more mixed. Several studies have reported positive associations between spatial ability and design performance, particularly when domain-specific spatial measures are used. Cho and Suh (2019) identified a moderately strong correlation between first-year design students' originality in an interior design project and their 2D-to-3D visualization scores from an architecture and interior design spatial ability test. However, no significant associations were found between students' creative design performance and their domain-general mental rotation and spatial visualization tasks performance. Similarly, Suh and Cho (2020) reported that third-year interior design students' domain-specific mental rotation and spatial visualization scores—but not the domain-general spatial scores—correlated moderately strongly with their creative design performance on an interior design project that emphasized spatial exploration. Similarly, Sutton and Williams (2010) found that first-year design students' graphics course grades significantly correlated with seven out of twelve sub-tests from a 3D spatial ability test, with small to medium effect sizes.

In contrast, other studies have found no significant correlation between creative design performance and spatial ability. For instance, Allen (2010) found no significant association between spatial visualization scores and creativity ratings of university students' interior design works. Similarly, Cho (2017) reported that first-year university design students' spatial visualization, mental rotation scores, or architecture-specific spatial ability did not meaningfully correlate with their design studio performance. Along the same lines, Reid and Sorby (2023) observed no significant correlation between university engineering students' creative design scores and their performance on a series of spatial tests, including mental rotation, mental cutting, surface development, and paper folding.

To sum up, existing findings do not point to a single, consistent explanatory pattern, indicating a complex and context-dependent relationship between spatial ability and creative design performance. Inconsistencies across these studies may partly reflect differences in spatial test characteristics (e.g., domain-specific versus domain-general) and approaches to assessing creative design quality (e.g., course grades versus task-based evaluations; criteria used for rating creativity). Moreover, the majority of existing research has focused on university-level students. Taken together, these gaps highlight the need for further investigation into how spatial ability may support creative design performance among school-aged pupils, especially in light of the growing emphasis on design-based practices in K-12 STEM education.

5.2.3 The roles of spatial ability and creativity in design performance

The relationship between spatial ability and creativity remains complex and not fully understood. Spatial ability is considered important for creativity, particularly in supporting the creation of innovative products and knowledge (Coxon, 2012; Kell et al., 2013; Newcombe & Shipley, 2015). There has been some evidence linking spatial ability to creative imagery. Burton and Fogarty (2003) found moderately strong correlations between various spatial ability tasks and creative imagery tasks developed by Finke et al. (1989). Their factor analysis revealed that three aspects of visual imagery—imagery quality, imagery speed, and self-report imagery—formed distinct yet structurally related factors within the spatial ability domain. Despite these overlaps, findings regarding the direct relationship between spatial ability and test-based creativity remain mixed (Cho, 2017; Morrison & Wallace, 2001), particularly among school-aged pupils (Antonietti & Colombo, 2011), indicating a need for further investigation.

Part of the ambiguity lies in whether creativity and spatial thinking involve the same kind of mental imagery. Several studies differentiated object-based imagery, which involves decoding object properties such as color, shape, and texture, from spatial imagery, which focuses on interpreting spatial relations, locations, and imagining spatial transformation (Borst & Kosslyn, 2010; Kozhevnikov et al., 2005; Kozhevnikov et al., 2013). These differences result in distinct visualization strategies (Kozhevnikov et al., 2005), with object-based imagery relating to artistic creativity and spatial imagery relating to scientific creativity (Kozhevnikov et al., 2013). Following this line of reasoning, spatial ability—known to reliably predict STEM performance (Uttal et al., 2013; Wai et al., 2009)—may have limited predictive power for tasks that emphasize artistic creativity than for scientific creativity. Design appears to sit at the intersection of these two types of creativity. On one hand, design methodologies position design as a structured, model-based process informed by computational principles (Gero, 1990; Rosenman & Gero, 2013). On the other hand, design has also been closely associated with the visual arts and approached through a more artistically creative lens (de la Harpe et al., 2009; Lindström, 2006). These seemingly contrasting perspectives underscore the importance of examining not only whether spatial ability contributes to creative design performance, but also whether this contribution is direct or shared through creativity-related processes such as divergent thinking or creative imagery.

5.2.4 Overview of the current study

The main goal of this study was to examine the relationship between spatial ability, creativity, and design performance in school-aged pupils, with a particular focus on creative design. Previous studies examining the relationship between spatial ability and creative design, such as those involving computer-aided design (Chang, 2014) and interior design projects (Suh & Cho, 2020), have not always clarified whether their design prompts explicitly emphasized creativity. Yet, psychological research on creativity underscores the importance of cognitive alignment between the task prompt and the targeted response (Nusbaum et al., 2014). Therefore, to ensure this alignment, this study employs a design task intended to encourage divergent thinking and the development of a creative prototype. Specifically, we focus on two early and critical stages of design: design ideation, which involves generating preliminary ideas or concepts that may later be developed or discarded (Lawson, 2005), and design prototyping, which advances these initial concepts by expressing their intended function, implementation, and aesthetic qualities in visual or tangible forms (Houde & Hill, 1997). In-depth analyses of designers' cognitive processes have identified these early phases in design as explicitly creative, heavily reliant on divergent thinking (Goldschmidt, 2016).

Prior research has not yet provided clear and direct evidence of how spatial ability and creativity independently or jointly relate to design performance among school-aged pupils. Given ongoing debates about the extent to which test-based creativity measures predict real-world creative performance, this study included a widely used measure of divergent thinking as an indicator of creative potential. Additionally, as imagery processes that lead to artistic creativity, such as vividness and expressiveness, are thought to differ from spatial imagery (Borst & Kosslyn, 2010; Kozhevnikov et al., 2005; Kozhevnikov et al., 2013), a measure of creative imagery was also included. This approach allowed us to examine whether imagination-focused creativity complements spatial-imagery-oriented constructs, such as spatial ability, in predicting pupils' creative performance in design. Pupils' spatial ability was assessed using the Spatial Reasoning Instrument (Ramful et al., 2017), which measures spatial skills considered critical for learning in upper primary and secondary education.

Specifically, the present study aimed to (1) investigate whether spatial ability and creativity (divergent thinking and creative imagery) correlate with pupils' creative design performance, (2) examine the unique contribution of spatial ability and creativity in predicting pupils' creative design performance, and (3) explore whether the relationship between spatial ability and creative design performance is direct or mediated through these creativity measures.

5.3 Methods

5.3.1 Participants

This study was conducted across thirteen classes from four international schools in the Netherlands, selected based on the expressed interest of the Design & Technology teachers and class availability. Participant attrition occurred due to pupils' personal or sick leaves and occasional session cancellations due to other school priorities. An additional nine pupils were excluded from the analysis due to limited English proficiency. The final sample consists of 201 pupils ($M_{\text{age}} = 12.52$, $SD_{\text{age}} = 1.14$, 108 girls, 93 boys). These pupils come from 33 different nationalities and had an average of 5.18 years ($SD = 2.80$) of experience studying in international schools where English is the primary language of instruction. This study has been approved by the Human Research Ethics Committee at TU Delft. Informed consent for participation and data use was obtained from the parents or legal guardians of all pupils included in the final sample.

5.3.2 Procedure

To accommodate differences in school schedules, three schools completed the above-mentioned tasks in four 45-minute sessions (sequence 1; n = 120), whereas one school completed the same tasks in three 60-minute sessions (sequence 2; n = 81). Details of the measures used can be found in the following section. A between-sequence comparability analysis was conducted to determine whether data from the two task sequences could be analyzed jointly.

Table 5.1 Tasks implemented in two different sequences

	Sequence 1 (n = 120) 45 minutes per session	Sequence 2 (n = 81) 60 minutes per session
Session 1	Design Task	Design Task Drawing Self-Efficacy
Session 2	Spatial Reasoning Instrument Drawing Self-Efficacy	Spatial Reasoning Instrument Verbal Ability Task
Session 3	Creative Imagery Task	Divergent Thinking Task Creative Imagery Task
Session 4	Divergent Thinking Task Verbal Ability Task	

5.3.3 Materials

Creative problem solving using a design task

The SnackSafe on the Beach design brief used in this study was adapted from a design problem that has been developed and tested at the Science Hub TU Delft. Pupils were tasked with developing a creative design enabling people to enjoy their fries on the beach, without worrying about seagulls stealing their fries. Pupils were told that the design should not harm the seagulls, should be easy to use on the beach, and that they were free to choose any target users. The open-ended nature of the problem allowed pupils to seek solutions from different angles, satisfying the need for using a relevant, ambiguous, and ill-defined problem to elicit non-routine thinking and creativity (Casakin, 2008; Custer, 1995).

Pupils individually documented their design processes—including problem exploration, idea generation, and concept development—in an A3-size foldable design booklet, inspired by the TERU project portfolio (Kimbell & Stables, 2007). Each design session began with a five-minute energizer highlighting the value of generating multiple ideas beyond initial thoughts. Pupils were instructed to first generate as many as six different ideas, followed by producing four other unusual and original ideas. These could be entirely new ideas, or improvements on and combinations of previous ideas. They were again encouraged to think from different viewpoints and try out various directions. These instructions were adapted from Butler (1987) and the YourTurn Design Tool (Klapwijk et al., 2021), aiming to clarify the goal of the task, sustain pupils' engagement, and encourage divergent thinking.

After ideation, pupils individually selected their most original idea for solving the design challenge and created detailed, annotated sketches of their design prototypes. All pupils received equivalent instructions and were given 40 minutes to complete both ideation and prototype sketching using only pencils. The task administration mirrored typical design classroom practices that facilitate idea development. Full details of the design task are available in Chapter 4.

Comparative assessment of the design task

To assess creativity seen in pupils' design ideation and design prototypes, we used comparative judgement, which holistically evaluates the relative quality of pupils' work through pairwise comparisons made by expert judges (Buckley, 2024a; Hartell & Buckley, 2021). This method relies on judges' collective expertise, making it particularly suitable for assessing complex, context-dependent constructs like creativity (Jones & Alcock, 2014), which may not be fully captured through rubric-based assessment (Sadler, 2008). Comparative judgement yields valid, reliable assessments and is intuitive to use, requiring only brief rater training (Pollitt, 2012; Verhavert et al., 2019). It has been used to assess work across educational levels and subjects (Bartholomew & Yoshikawa-Ruesch, 2018), including English essays (Steedle & Ferrara, 2016), math problem solving (Jones & Inglis, 2015), and Design and Technology projects (Buckley et al., 2022). Given previous findings indicating that creativity was not among the most frequently considered criteria if judges were to evaluate the overall quality of design works (Bartholomew et al., 2020; Buckley et al., 2022), we specifically highlighted the goal of evaluating creativity during both the pre-assessment training and the evaluation itself.

Full details of the comparative judgement analysis can be found in Chapter 4. Pupils' design ideation and prototype scores were calculated on the Adaptive Judgement App (Buckley, 2024a, 2024b), through fitting the Bradley-Terry-Luce model using the Supplementary Item Response Theory Models (sirt) package (Robitzsch, 2021) in R. The scale separation reliability, which is an indicator of inter-rater reliability in comparative judgement, was 0.834 for design ideation and 0.753 for design prototype, reflecting good reliability that is consistent with previous studies (Bramley & Vitello, 2019; Jones & Alcock, 2014; Verhavert et al., 2019). See below for examples of one of the top- and bottom-ranked outcomes in ideation and prototyping from four different pupils. Annotations that were difficult to read have been transcribed to facilitate judges' understanding.

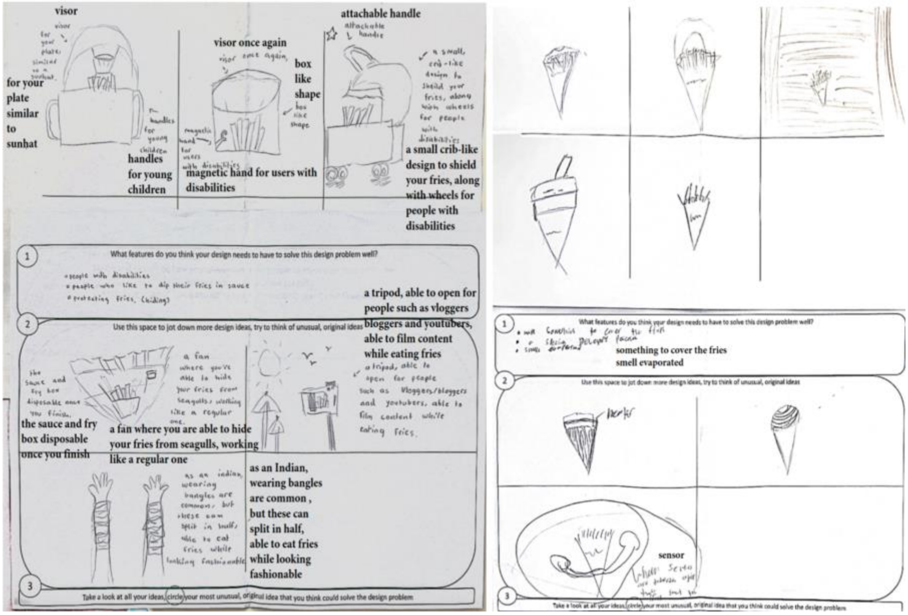


Figure 5.1 One of the top-ranked (left) and bottom-ranked (right) ideation outcomes based on comparative judgement results

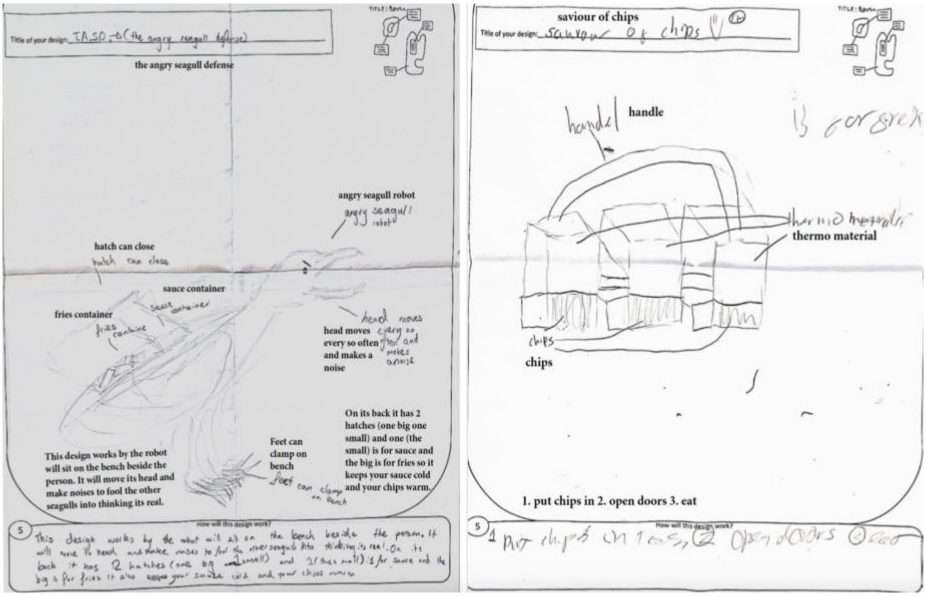


Figure 5.2 One of the top-ranked (left) and bottom-ranked (right) design prototypes based on comparative judgement results

Divergent thinking task

The Alternate Uses Tasks, one of the most widely used divergent thinking measures (Saretzki et al., 2024), was used to assess divergent thinking. Pupils were asked to respond individually on a Qualtrics form either on a laptop or a tablet, whichever device they used daily for school. They first typed out all the unusual and original uses of a brick as they could think of in 3 minutes in one textbox. The instruction highlighted the need to be creative as recommended by Nusbaum et al. (2014). The pupils then retyped the two uses in another textbox that they considered to be the most creative out of all the uses they thought of. They repeated this process for another prompt on unusual uses of a paperclip. Three raters, who were design students with experience working with school-aged pupils, received a 30-minute training on holistically rating the entire set of responses from each pupil on a five-point scale ranging from 1 (not creative) to 5 (very creative) while considering the following desired qualities: uncommonness, remoteness, and cleverness (Silvia et al., 2008). Raters were presented with pupils' responses in their original form without typo correction. To prevent judges' fatigue, we asked judges to carefully plan the number of responses they rate at one time and use scoring heuristics[5] recommended by Forthmann et al. (2017) when a decision is hard to make.

The intraclass correlation coefficient (ICC) among three raters based on absolute agreement was ICC(3,3) = 0.733, 95% CI [0.654, 0.795] for grading brick, and ICC(3,3) = 0.728, 95% CI [0.656, 0.787] for grading paperclip, both indicating moderate inter-rater reliability. These values are comparable to previous research reporting raters using the snapshot scoring method to rate Alternate Uses Tasks (e.g., Benedek et al., 2014; Forthmann et al., 2016; Forthmann et al., 2017).

Creative imagery task

To measure creative mental imagery, we used the Test of Creative Imagery Abilities (TCIA) (Jankowska & Karwowski, 2015, 2020), an instrument developed for educational contexts and particularly suited to evaluating young learners' creative imagery abilities (Dziedziewicz & Karwowski, 2015). The TCIA measures three dimensions: imagery vividness (the ability to generate clear and elaborated visualization of mental images), imagery originality (the ability to generate novel and unique mental images), and imagery transformation (the ability to actively manipulate, modify, and transform the mental images generated). The task was administered in paper-and-pencil format following the TCIA test booklet (Jankowska & Karwowski, 2020). Each pupil individually responded to seven different initial figures by first generating as many mental images as possible based on the given initial figure. They then selected their most unique mental image and created an elaborate drawing and a brief description. The instruction emphasized the need to change and transform their mental image and the need to develop original and creative drawings. Pupils were given 45 minutes of class time to finish the seven tasks, aligning

[5] Two scoring heuristics to consider: (1) raters may compare and provide compensation between a participant's typical performance perceived from all the given responses and his/her most creative response; in other words, a strong standout idea might elevate the perceived creativity of a participant's full set of responses; (2) raters may consider scoring a set of responses that involve ideas from different semantic categories higher than a set that only involve ideas from one semantic category (Forthmann et al., 2017, p. 137).

with the recommended task time.

Six design students with experience working with school-aged pupils received 30 minutes of training on the TCIA scoring manual. All raters practiced rating five sample drawings during training and clarified any questions they had with the researcher. Two raters were assigned to each scoring dimension: vividness, originality, and transformation. Each drawing was rated on a 3-point scale (0, 1, or 2), and each rater completed a total of 1407 ratings. For vividness, the inter-rater reliability based on absolute agreement was ICC(3,2) = 0.816, 95% CI [0.796, 0.834], indicating good reliability (Koo & Li, 2016). For originality, the inter-rater reliability was ICC(3,2) = 0.798, 95% CI [0.756, 0.831], indicating good reliability. For transformation, the inter-rater reliability was ICC(3,2) = 0.684, 95% CI [0.632, 0.727], indicating moderate reliability.

Spatial reasoning task

We assessed spatial ability using the Spatial Reasoning Instrument (SRI), which was developed to address the multidimensional construct of spatial ability and has been validated with upper primary and secondary school pupils (Lowrie et al., 2019; Ramful et al., 2017). The SRI contains ten items from each of the three sub-dimensions: mental rotation, spatial orientation, and spatial visualization (see sample items in Figure 5.3). This measure correlates well with established spatial ability measures (Ramful et al., 2017) and targets both static and dynamic spatial transformations. An additional feature that distinguishes this instrument from other spatial ability measures—such as the Card Rotation Test or the Paper Folding Test (Ekstrom et al., 1976)—is its use of contextual representations of spatial problems that are familiar to pupils while simultaneously challenging them to apply spatial thinking across a range of scenarios.

Pupils completed this 30-item instrument individually in paper-and-pencil format within 45 minutes. Because the SRI prioritizes accuracy over speed to accommodate pupils' developing spatial understandings, the allotted time is sufficient for most learners (Ramful et al., 2017). The internal consistency reliability of the 30-item spatial ability measure was assessed using the Kuder-Richardson Formula 20 (KR-20), an analog to Cronbach's alpha since we scored this test dichotomously (1 = correct, 0 = incorrect). The measure demonstrated good internal consistency (KR-20 = 0.804), comparable to the Cronbach's alpha of 0.849 reported by Ramful et al. (2017), indicating that the items coherently assess the proposed spatial ability construct.

Verbal ability task

Peabody Picture Vocabulary Test (PPVT) Fourth Edition in English, a verbal vocabulary comprehension test originally developed by Dunn and Dunn (2007), was included as a measure to control for possible confounding effects of verbal ability. Previous research indicated that verbal ability correlates with spatial ability (Casey et al., 2015) and predicts verbal divergent thinking measured by Alternate Uses Tasks (Silvia et al., 2013; Forthmann et al., 2019). Meanwhile, verbal abilities may also influence how well individuals comprehend the design brief or articulate their designs (Tomes et al., 1998). This measure was administered digitally without imposing a time limit. On laptops or tablets, pupils identified a pictorial representation of the word heard from their individual headphones. All pupils completed the test within 20 minutes. Scores were generated through the testing service website as individual score summary reports, which were standardized based on the age norm using a 95% confidence interval.

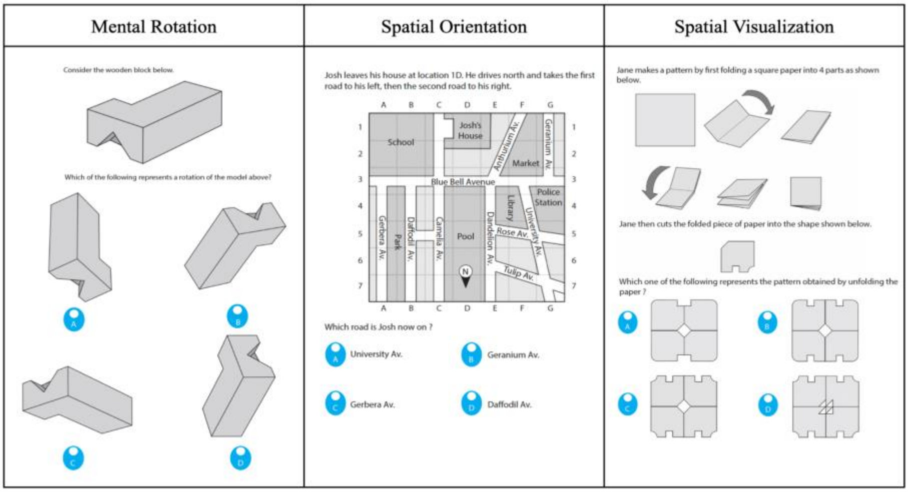


Figure 5.3 Examples of questions from the three dimensions in the SRI (Ramful et al., 2017)

Drawing Self-Efficacy Instrument

Much evidence has suggested that drawing plays a key role in both expert and novice designers' design problem-solving (Ferguson, 1977; Suwa et al., 1998; Tedjosaputro et al., 2018). Beyond communicating design ideas, drawings evolve with the development and transformation of mental imagery of ideas (Tedjosaputro et al., 2018) and facilitate the creative re-interpretation and re-representation of (spatial) information from ideas (Suwa et al., 1998). To take into consideration that pupils' perceived ability to draw may have influenced their design and creative measure outcomes, we explored whether pupils' self-reported efficacy in drawing, measured by the Drawing Self-Efficacy Instrument (DSEI) (Jaison et al., 2021), relates to other variables investigated above. This instrument, surveying individuals' abilities in drawing to communicate or create ideas and drawing specific objects, has been validated by Jaison et al. in high school art and engineering education contexts ($\alpha = 0.943$). From 0 (lowest) to 10 (highest), pupils self-reported how confident they felt about performing tasks mentioned in 13 different statements, such as "Drawing to communicate ideas to others," "Drawing a 3D object," and "Drawing something from my imagination."

5.4 Results

5.4.1 Main analysis

We first conducted descriptive analysis and tested for normality. Since the Shapiro-Wilk tests revealed non-normal distributions for most of our variables, we employed nonparametric tests for subsequent analysis in R Studio 4.3.2 and used bootstrapping where nonparametric alternatives were not suitable. We then performed a Mann-Whitney U test to decide whether the data collected from two different task sequences (see Table 5.1) were similar enough to be analyzed together. Besides differences in age (with pupils following sequence two being slightly older than those following sequence one), no significant differences were observed for their design scores, creativity measures, spatial ability, or verbal ability. As age was not the focus of our analysis, we considered it reasonable to treat these comparable data as a single dataset for subsequent analysis.

A Spearman correlation analysis was then performed to examine the relationships between measures. To better understand how different measures predict design performance and to determine whether spatial ability predicts design performance beyond other measures, we conducted a hierarchical regression analysis. Mediation analysis was then performed to verify whether the effect of spatial ability on design performance is direct or operates indirectly through a mediator influencing design scores.

Correlation analyses

Spearman's rank correlation was used to examine the monotonic relationships between the variables because many of the variables showed significant deviations from normality. Bootstrapping was used to calculate 95% confidence intervals for the correlation estimates, with 5,000 resamples to enhance the stability and robustness of the results, given the non-normal distribution of the data. To control for the False Discovery Rate (FDR) given the multiple correlation tests conducted, we applied the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) using a common threshold of $q = .05$. After FDR correction, the correlations between divergent thinking and design prototype ($r = 0.151$, $p = .032$, 95% CI [0.013, 0.288]), between verbal ability and divergent thinking ($r = .144$, $p = .041$, 95% CI [-0.001, 0.285]), and between verbal ability and creative imagery abilities ($r = 0.153$, $p = .030$, 95% CI [0.004, 0.290]) were no longer statistically significant.

Hierarchical regression analyses

Hierarchical regression analyses were conducted in R Studio 4.3.2 to systematically examine the contribution of different predictors to design prototype and design ideation performance, while controlling for age and verbal ability. To account for differences in measurement scales across predictors, all variables were standardized to z-scores prior to analysis. Age and verbal ability were entered as covariates at the first step to control for their potential effects on design performance.

Table 5.2 Descriptive statistics for raw scores and correlation matrix after bootstrapping

	M	SD	Design ideation	Design prototype	Divergent Thinking	Creative imagery abilities	Spatial ability	Verbal ability	Age
Design ideation	0.00	1.62	1						
Design prototype	0.00	1.12	0.369** [0.240, 0.490]	1					
Divergent thinking	6.14	1.29	0.201** [0.063, 0.332]	0.151* [0.013, 0.288]	1				
Creative imagery abilities	16.30	6.11	0.294** [0.164, 0.413]	0.181* [0.042, 0.317]	0.253** [0.115, 0.380]	1			
Spatial ability	20.30	4.97	0.165* [0.022, 0.310]	0.234** [0.094, 0.372]	0.197** [0.065, 0.332]	0.098 [-0.051, 0.241]	1		
Verbal ability	105.23	12.18	0.127 [-0.011, 0.268]	0.094 [-0.044, 0.229]	0.144* [-0.001, 0.285]	0.153* [0.004, 0.290]	0.384** [0.254, 0.503]	1	
Age	12.52	1.14	-0.016 [-0.157, 0.128]	0.076 [-0.059, 0.209]	-0.126 [-0.261, 0.012]	0.013 [-0.125, 0.149]	0.137 [-0.004, 0.263]	0.017 [-0.125, 0.161]	1

Note. *$p < .05$ (two-tailed). **$p < .01$ (two-tailed). Values in brackets represent 95% confidence intervals

Table 5.3 Hierarchical regression analyses predicting design ideation scores

Predictor	Step 1 (β, 95% CI)	Step 2 (β, 95% CI)	Step 3 (β, 95% CI)	Step 4 (β, 95% CI)
Age	-0.050 [-0.189, 0.089]	-0.025 [-0.163, 0.113]	-0.032 [-0.166, 0.103]	-0.041 [-0.178, 0.096]
Verbal ability	0.109 [-0.030, 0.248]	0.082 [-0.056, 0.220]	0.056 [-0.079, 0.192]	0.037 [-0.107, 0.181]
Divergent thinking	—	0.199** [0.060, 0.338]	0.144* [0.005, 0.284]	0.134 [-0.008, 0.276]
Creative imagery abilities	—	—	0.234* [0.096, 0.373]	0.236** [0.097, 0.374]
Spatial ability	—	—	—	0.059 [-0.087, 0.206]
R^2	0.014	0.052	0.103	0.106
ΔR^2	—	0.038	0.051	0.003
F	1.41	3.62*	5.64***	4.63***

Note. **p* < .05 (two-tailed). ***p* < .01 (two-tailed). ****p* < .001 (two-tailed). Standardized regression coefficients (β) are reported. Values in brackets represent 95% confidence intervals.

Table 5.4 Hierarchical regression analyses predicting design prototype scores

Predictor	Step 1 (β, 95% CI)	Step 2 (β, 95% CI)	Step 3 (β, 95% CI)	Step 4 (β, 95% CI)
Age	0.068 [-0.071, 0.206]	0.086 [-0.053, 0.224]	0.081 [-0.056, 0.218]	0.048 [-0.089, 0.185]
Verbal ability	0.116 [-0.023, 0.255]	0.097 [-0.043, 0.236]	0.077 [-0.061, 0.215]	0.010 [-0.134, 0.153]
Divergent thinking	—	0.145* [0.005, 0.285]	0.103 [-0.039, 0.245]	0.067 [-0.075, 0.209]
Creative imagery abilities	—	—	0.181* [0.040, 0.322]	0.185** [0.046, 0.324]
Spatial ability	—	—	—	0.207** [0.060, 0.353]
R^2	0.019	0.039	0.069	0.105
ΔR^2	—	0.020	0.030	0.036
F	1.87	2.65*	3.65**	4.57***

Note. **p* < .05 (two-tailed). ***p* < .01 (two-tailed). ****p* < .001 (two-tailed). Standardized regression coefficients (β) are reported. Values in brackets represent 95% confidence intervals.

Age, verbal ability, divergent thinking, creative imagery abilities, and spatial ability together significantly predicted design ideation scores, $F(5, 195) = 4.63$, $p < .001$, accounting for 10.6% of the variance ($R^2 = 0.106$, adjusted $R^2 = 0.083$), and design prototype scores, $F(5, 195) = 4.57$, $p < .001$, accounting for 10.5% of the variance ($R^2 = 0.105$, adjusted $R^2 = 0.082$). Of all the predictors, only creative imagery ($\beta = 0.236$) significantly predicted design ideation scores in the final model. The addition of spatial ability contributed very little to explaining design ideation scores. Both creative imagery ($\beta = 0.185$) and spatial ability ($\beta = 0.207$) significantly predicted design prototype scores. Notably, spatial ability emerged as the strongest predictor, uniquely explaining an additional 3.6% of the variance in design prototype scores beyond other predictors. When entered at Step 2, divergent thinking contributed significantly to both design ideation scores and design prototype scores. However, when all other variables have been entered, divergent thinking was no longer a significant contributor to either ideation scores ($\beta = 0.134$, $p = .064$) or prototype scores ($\beta = 0.067$, $p = .354$).

The Shapiro-Wilk test was conducted to assess the normality of the residuals in our final regression models. The residuals were found to be normally distributed both for the model predicting design ideation scores (Shapiro-Wilk $W = 0.994$, $p = .622$) and the model predicting design prototype scores (Shapiro-Wilk $W = 0.987$, $p = .058$). Both models meet the assumption of homoscedasticity, as indicated by the Breusch-Pagan test results: $BP = 2.203$, $p = .820$ for the design ideation model, and $BP = 4.06$, $p = .541$ for the design prototype model.

Multilevel modeling results controlling for school-level variance

To account for potential clustering by school, we fitted multi-level models that treated school as a random intercept. The model was fitted using Restricted Maximum Likelihood estimation (Pinheiro & Bates, 2000), which aimed to provide more robust estimates of variance components with modest group sizes (four schools). Fixed-effect estimates did not deviate much from those obtained in the hierarchical regressions (Tables 5.3 & 5.4). At the group level, school accounted for 6.1% in design ideation scores and 11.6% of the variance in design prototype scores, indicating that most of the variability occurred at the individual pupil level.

Mediation analyses

Given the significant correlations between spatial ability, divergent thinking, and design performance (Table 5.2), we conducted mediation analyses with 5,000 nonparametric bootstrap resamples to investigate whether divergent thinking mediated the effect of spatial ability on design ideation and design prototype performance.

For design ideation, the average direct effect of spatial ability was 0.059 ($p = .422$, 95% CI [-0.092, 0.210]), and the total effect was 0.085 ($p = .254$, 95% CI [-0.062, 0.230])—neither of which was statistically significant. The average causal mediation effect (ACME) was 0.026, which was not statistically significant either ($p = .059$, 95% CI [-0.001, 0.060]). Although the proportion mediated was 30.25%, its wide confidence interval ($p = .290$, 95% CI [-2.401, 3.050]) suggests considerable uncertainty about the magnitude and direction of the mediation effect.

For design prototype, the analysis revealed a significant average direct effect of 0.207 ($p = .008$, 95% CI [0.059, 0.360]) and a total effect of 0.220 ($p = .004$, 95% CI [0.072, 0.370]). The ACME of 0.013 ($p = .363$, 95% CI [-0.015, 0.050]) and the proportion mediated (5.86%; p = .364, 95% CI [-0.090, 0.300]) were both small and not statistically significant, indicating that divergent thinking did not meaningfully mediate the relationship between spatial ability and design prototype scores.

For completeness, we also conducted mediation analyses with creative imagery as a mediator. For design ideation, the analysis revealed a non-significant average direct effect of 0.059 (p = .400, 95% CI [-0.089, 0.210]) and total effect of 0.076 (p = .320, 95% CI [-0.082, 0.230]). The ACME (0.016; p = .340, 95% CI [-0.018, 0.060]) and proportion mediated (21.69%; p = .440, 95% CI [-1.742, 2.310]) also failed to reach significance. For design prototype, the average direct effect of 0.207 (p = .008, 95% CI [0.053, 0.360]) and total effect of 0.220 (p = .006, 95% CI [0.064, 0.380]) were both significant. However, the ACME (0.013; p = .333, 95% CI [-0.013, 0.050]) and the proportion mediated (5.88%; p = .334, 95% CI [-0.084, 0.280]) were not significant, indicating that creative imagery did not meaningfully mediate the relationship between spatial ability and design prototype scores.

5.4.2 Exploratory analysis

Correlation matrix for spatial ability sub-scores

We further explored how the three sub-scores of the SRI, mental rotation (MR), spatial orientation (SO), and spatial visualization (SV), correlate with design ideation and design prototype scores separately, using Spearman correlation along with bootstrapping based on 5,000 resamples to calculate 95% confidence intervals for the correlation estimates (Table 5.5). All three sub-scores of the SRI correlated with design prototype scores. Only spatial orientation correlated significantly with design ideation scores, although its effect sizes did not differ much from that of mental rotation or spatial visualization.

Table 5.5 Spearman correlation between design scores and spatial sub-scores

	Design ideation	Design prototype	SRI_MR	SRI_SO	SRI_SV
Design ideation	1				
Design prototype	0.369** [0.241, 0.490]	1			
SRI_MR	0.124 [-0.018, 0.266]	0.185* [0.049, 0.324]	1		
SRI_SO	0.173* [0.035, 0.309]	0.188** [0.045, 0.327]	0.429** [0.301, 0.542]	1	
SRI_SV	0.134 [-0.008, 0.261]	0.195** [0.057, 0.334]	0.552** [0.443, 0.647]	0.320** [0.187, 0.439]	1

Note. *p < .05 (two-tailed). **p < .01 (two-tailed). Values in brackets represent 95% confidence intervals.

Correlation with drawing self-efficacy

Less than half of the pupils included in our main analysis were able to complete this instrument (n = 82). The rest were unable to submit the self-report either due to a loss of class time caused by unexpected interruptions or delays before or during class, or due to technical difficulties with their devices. A Spearman correlation test revealed no significant correlations between pupils' reported drawing self-efficacy and scores in design, creativity measures, spatial ability, or verbal ability. The only correlation statistic close to significance was between drawing self-efficacy and pupils' age (r = -0.216, p = .051, 95% CI [-0.419, 0.009]).

5.5 Discussion

This study examined the unique and combined contributions of divergent thinking, creative imagery, and spatial ability to creative design performance among primary and secondary school pupils. Divergent thinking, creative imagery, and spatial ability all correlated significantly with creative design performance. Design ideation and design prototype scores were only moderately correlated, suggesting that while performance in these two phases of design were related, they may draw on the predictor variables in different ways. Creative imagery was a consistent and significant predictor of both design ideation and prototype scores. However, the predictive value of divergent thinking diminished once creative imagery and spatial ability were included in the regression models. Spatial ability uniquely predicted design prototype performance, accounting for a small but significant proportion of variance beyond what was explained by creativity measures. In contrast, its relationship with design ideation was not statistically significant when other predictors were accounted for, implying that spatial ability may play a more prominent role in the prototyping phase of design than in idea generation. The following sections delve into the unique and shared contributions of creativity and spatial ability to pupils' creative design performance.

5.5.1 Creativity measures and creative design performance

Our findings revealed a nuanced relationship between the two creativity measures and pupils' creative design performance. Divergent thinking did not significantly correlate with design prototype scores after controlling for the false discovery rate. Furthermore, its initially significant contribution to both design ideation and prototype performance diminished once creative imagery and spatial ability were added to the regression model. This suggests that divergent thinking did not uniquely predict creative performance in our selected design task, or that its effect may be shared with creative imagery and spatial ability. Although divergent thinking has long been regarded as a meaningful estimate of creative potential and reliably predicts real-world creative achievement (Jauk et al., 2014; Kim, 2008; Plucker, 1999; Runco & Acar, 2012), our findings can be understood in light of Runco's (2010) argument that measures such as divergent thinking tasks provide only an "estimate of the potential for creative thinking and problem solving" (p. 424), and do not guarantee that such ideational performance will generalize across domains or translate into real-world creative outcomes. In a similar line of inquiry, Cho (2017) found no significant correlation between another measure of divergent thinking (TTCT) and freshmen's design studio performance. Together, our findings and those of Cho (2017) highlight a potential gap between domain-general assessments of divergent thinking and the specific types of creativity required for design.

Creative imagery correlated significantly with both design ideation and design prototype, with its association to ideation approaching a medium effect size. In the regression analysis, creative imagery abilities remained a significant predictor for both design ideation and prototype scores after controlling for all other variables. These findings aligned with earlier qualitative (Athavankar, 1997) and quantitative (Kokotovich & Purcell, 2000) research, adding to the evidence that visual imagery plays a direct and meaningful role in design. For instance, creative mental imagery may facilitate the rapid generation of diverse design ideas (Tedjosaputro et al., 2018), as well as enhance the extent to which a design prototype coherently integrates the various elements that need be addressed (Christiaans, 2002).

5.5.2 Spatial ability and design ideation scores

Spatial ability showed a small but statistically significant correlation with design ideation scores in correlation analysis, suggesting that pupils with higher spatial ability generated more creative design ideas for this task. However, spatial ability did not significantly predict design ideation scores once creativity measures were accounted for. In other words, having good spatial skills did not provide much additional benefit for design ideation beyond what was already explained by creativity. This may be due in part to the nature of the design task: the ideation phase emphasized generating a wide range of uncommon ideas, with quality judged primarily on variety and novelty rather than on functionality or feasibility (see Chapter 4). It is expected that spatial ability may be more important for the latter (Raju et al., 2023).

Through exploratory analysis, we noticed that only spatial orientation, but not mental rotation or spatial visualization, significantly correlated with design ideation scores. We speculate that this might be because the design task itself requires not only thinking about the fries packages but also about the space people are in. As reported in Chapter 4, more than half of the solutions designed were beyond packaging design, with a number of them involving distracting seagulls or directing them to certain places that may involve considerations of spatial orientation and spatial arrangement. Therefore, to come up with creative ideas for this specific design brief, it might have been helpful for students to also consider the overall beach environment and orient themselves around their products. That said, as the effect sizes across the three spatial subskills were comparable, this speculation should not be interpreted as evidence that spatial orientation played a more important role.

5.5.3 Spatial ability and design prototype scores

Spatial ability measured by the SRI—and each of its three dimensions—showed small correlations with design prototype scores, echoing prior findings that spatial skills are linked to creative design performance (Chang, 2014; Suh & Cho, 2020).

In the regression analysis, spatial ability significantly predicted and uniquely explained 3.6% of the variance in design prototype scores after controlling for pupils' age, verbal ability, divergent thinking, and creative imagery. Moreover, spatial ability's contribution to design prototype scores was not mediated by divergent thinking or creative imagery, suggesting that spatial ability has played a direct role in predicting creative design performance rather than through creativity. The unique contribution of spatial ability to STEM performance, as reported in previous studies, appears to fall within a comparable range. For example, Hodgkiss and colleagues (2018) found that mental folding and spatial scaling, respectively, accounted for 4% and 2% of the variance in 7-to-11-year pupils' physics performance. Träff et al. (2019) reported that mental rotation assessed in 9-to-10-year-olds explained 4.6% of the variance in pupils' physics skills at 12-13 years. Due to notable differences in study aims, sample characteristics, and the total variance reported, we do not intend to draw direct comparisons between our findings and those of previous studies. Nevertheless, the similar magnitudes observed across studies provide support for the modest yet significant role spatial ability plays in predicting STEM performance.

Our finding that spatial ability emerged as the strongest predictor for design prototype scores beyond creativity measures may be better understood if we further look at what determined the quality of the design prototype. In the analyses examining the criteria used for assessing the design task in this study (see Chapter 4), it was reported that novelty was a key criterion (frequency = 0.257). However, usability—mainly reflecting functionality and user experience—was mentioned even more frequently (frequency = 0.315). These findings align with a broad consensus within and beyond the design community that creative design should be not only novel but also useful and appropriate (Cropley & Kaufman, 2012; Howard et al., 2008; Sternberg & Lubart, 1999). In our present study, it appears that pupils with higher spatial ability created designs that were not just novel but also useful and practical (see Figures 5.1 and 5.2 for examples) according to the expert judges. In Chang's (2014) study, while pupils' spatial ability correlated significantly with overall creative performance (r = 0.351) and their subscales—novelty, aesthetics, and functionality—the correlation between spatial ability and functionality was the strongest (r = 0.401). In addition, Raju and colleagues (2023) found that first-year college students with high spatial ability spent almost double the time focusing on analyzing the workability of their design as compared to those with low spatial ability. Taken together, this line of findings lends to the speculation that spatial ability is not only needed for generating novel ideas, but also for turning those ideas into functional, workable designs. Spatial ability may contribute to functionally creative solutions by enabling easier 2D-to-3D visualization (Cho & Suh, 2019), or by supporting multistep mechanical reasoning (Hegarty & Sims, 1994), which is needed to arrange different design components into a functioning system. Future research is needed to examine how spatial ability is engaged in functionally creative aspects of designs during prototyping, as well as testing and iteration phases.

It is also worth noting that our findings differ from previous studies reporting that creative design performance among university design students was associated only with domain-specific spatial ability but not with domain-general measures (Cho & Suh, 2019; Suh & Cho, 2020). In the current study, domain-general spatial ability, as measured by the SRI, correlated with pupils' creative design performance, consistent with Chang's (2014) findings among high school students. One possible explanation for this difference lies in learners' stage of expertise. As suggested by Hambrick et al. (2012), domain-general spatial skills may be stronger predictors of academic performance in the early stages of learning. More advanced learners likely rely more on domain-specific knowledge than on general cognitive abilities such as spatial reasoning. From this perspective, and in light of suggestions that spatial ability may act as a 'gatekeeper' in early STEM learning (Uttal & Cohen, 2012), it is worthwhile to further explore how spatial skills development can be supported among young and novice design learners.

5.5.4 Spatial ability and creativity measures

This study also examined whether spatial ability and creativity measures are meaningfully related or largely distinct. Our finding that spatial ability correlated significantly with divergent thinking, despite the small effect size, is particularly interesting, given that spatial ability tasks are often characterized as convergent and analytically oriented (Guignard & Lubart, 2007). This finding may relate to previous studies identifying a small correlation between divergent thinking and fluid intelligence, which was partly measured by tasks that demand spatial ability, such as paper-folding (Beaty et al., 2014; Nusbaum & Silvia, 2011). However, to our knowledge, few studies have directly examined the relationship between spatial ability and divergent thinking, leaving this connection underexplored in the literature. Chrysikou and Thompson-Schill (2011) found through fMRI data that generating creative uses for everyday objects was associated with increased activation in brain regions involved in visual processing. Building on this account, we speculate that even when divergent thinking is assessed through textual responses—such as pupils typing that a brick can be "cut into different pieces and make something with it like a statue"—visual-spatial thinking processes, such as mental transformation, may still occur in pupils' minds as they ideate unusual uses for a common item.

We anticipated that spatial ability and creative imagery scores would be related due to their seemingly shared mechanisms related to imagery (Carroll, 1993). However, the lack of significant correlation found between them in this current study is not entirely surprising, as others have argued for distinguishing the process of creating vivid and expressive object-focused visualizations that emphasize shape and colour from using imagery to process spatial relations and spatial transformations (Kozhevnikov et al., 2005). For instance, the TTCT-figural, which measures the creative potential to create original and expressive visualizations, was considered a better predictor for 'artistic creativity' rather than 'scientific creativity' (Kozhevnikov et al., 2013). In a line of inquiry similar to ours, Cho (2017) also found no significant correlation between freshmen's spatial scores and scores from their TTCT-figural. Taken together, the non-significant correlation between creative imagery abilities and spatial ability in this study may be attributed to the different visualization processes needed to complete these tasks, lending support to the conceptual distinction between spatial and object-based imagery processes (Kozhevnikov et al., 2005).

5.6 Limitations and future directions

Despite including verbal and figural creativity measures as well as spatial and verbal ability measures, our model accounted for around 10% of the variance in pupils' design ideation and prototype scores. The limited amount of explained variance may be attributed to the fact that the predictors selected in this study—spatial ability and creativity measures—were domain-general tests not specifically tailored to design contexts. Their broad scope may have limited their explanatory power for a context-specific design task (Palmiero et al., 2015), contributing to the small effect sizes observed in the correlation and regression analyses. Future research may benefit from incorporating a battery of spatial tests (Ormand et al., 2014), rather than a single composite measure, to tease apart which sub-skills best predict creative design performance. Besides, instead of relying on domain-general creative potential measures, future studies could consider using measures that better assess 'functional creativity' (Cropley & Kaufman, 2012), which more closely reflect qualities valued in design contexts.

A further limitation concerns the design task used to assess pupils' creative design performance. This study employed a single, time-constrained 45-minute design task, which did not allow for multiple cycles of iteration and refinement typically involved in long-term design projects. As a result, pupils' performance may have been influenced by time pressure and short-term cognitive demands, such as processing speed and working memory. Our findings should therefore be interpreted with caution when generalizing to other design contexts. Future research could benefit from employing multiple design tasks with varying time frames that reflect a range of contexts and scenarios to more comprehensively reflect pupils' design capability (Kimbell & Stables, 2007).

There are likely many other factors influencing the design processes and outcomes, such as knowledge base (Rosenman & Gero, 2013) and domain knowledge (Christensen & Ball, 2016), intrinsic motivation (Amabile & Gitomer, 1984), shifting between divergent and convergent thinking (Goldschmidt, 2016), attention allocation during problem-solving (Asghar et al., 2024), and crafting skills (Buckley et al., 2022). Future studies may want to explore whether these factors meaningfully mediate the relationship between spatial ability and creative design performance.

Finally, since participation was voluntary, our sample may have been subject to self-selection bias: teachers who offered their classes may be those who were confident in their pupils' abilities or particularly interested in design research, potentially limiting the generalizability of our findings. Future studies with larger samples and accounting for additional covariates may help clarify the strength and direction of these relationships and yield more stable effect-size estimates.

5.7 Conclusion

This study contributes new insights into how different cognitive abilities relate to creative design performance among school-aged pupils. Our findings suggest that pupils' performance in design ideation and design prototype drew on partially distinct cognitive skill sets. Specifically, while spatial ability did not significantly predict creativity in design ideation once creativity measures were accounted for, it emerged as a significant and unique predictor of design prototype quality, predicting variance above and beyond the contributions of both divergent thinking and creative imagery. To our knowledge, this provides some of the first evidence of a direct link between school-aged pupils' spatial ability and creative design performance—an area that has received limited attention to date. Findings from mediation analyses further revealed that spatial ability influenced design prototype quality through pathways independent of divergent thinking or creative imagery, pointing to a direct role of spatial ability in the visual and physical realization of ideas. We speculate that spatial ability may support the development of prototypes that are not only novel but also functional, possibly by enabling pupils to visualize how a design operates and how its form, function, and context can be woven into an original solution. Future research should explore these possibilities more fully by examining design performance during the testing and iteration phases and by investigating additional mediators. Placing these findings within a broader educational context and as design continues to take center stage in integrated STEM education, deepening our understanding of design's cognitive underpinnings remains a valuable and timely pursuit. While spatial ability is often regarded as a technical aptitude, the present study suggests that it could be an important driver of functional creativity in design.

CHAPTER 6
Conclusions, discussion, and future directions

6.1 Introduction

This dissertation presented a multi-method investigation into the forms of spatial thinking pupils engage in during design activities and how their spatial ability contributes to design outcomes. It began with a broad review of strategies for fostering spatial skills development in integrated STEM education. A considerable body of research has highlighted the important role spatial ability plays in STEM learning and career (Atit et al., 2020; Kell et al., 2013; Newcombe, 2017; Uttal & Cohen, 2012; Wai et al., 2009). By examining spatial thinking within K-12 design processes, this dissertation sought to bridge spatial ability research and design education, thereby contributing to a more holistic understanding of how spatial thinking supports innovation and engagement in STEM learning.

Across four studies, it became increasingly clear that spatial thinking plays a meaningful and multifaceted role in creative design and making. Chapter 2 showed how pupils constructed tangible visual–spatial representations of numerical information through hands-on data physicalization activities, thereby engaging in embodied forms of spatial thinking. Chapter 3 extended this focus to an open-ended design-by-analogy task, revealing how pupils used visual-spatial thinking to abstract biological features and explore diverse solutions. These findings suggested potential distinctions between spatial thinking in creative design contexts and that assessed by spatial tests. Building on these insights, Chapter 4 introduced comparative judgement as a method to evaluate the creativity and quality of pupils' design ideas and prototypes, demonstrating its potential as a holistic and nuanced approach to assessing creative design. Finally, Chapter 5 examined the unique contributions of spatial ability to creative design performance and found that, after accounting for age, verbal skills, and creativity measures, spatial ability significantly predicted design prototype quality. This finding underscored the particular importance of spatial ability in supporting pupils' translation of initial ideas into developed solutions. The following sections synthesize the key findings and conclusions from these four studies, followed by a general discussion, an examination of limitations, directions for future research, and practical implications for design education.

6.2 Findings and conclusions per study

In the context of data physicalization making (Chapter 2), pupils created a variety of visual-spatial representations to convey numerical information. The diverse materials available not only enabled open-ended approaches to constructing tangible visualizations but also offered immediate visual feedback as pupils experimented with different physicalization designs. Pupils' interactions with materials shaped both their representational choices—such as mapping quantities to sizes, proportionally unitizing materials and arranging them ordinally—and their embodied spatial reasoning, including conceptualizing, crafting, evaluating, and refining designs while anticipating viewers' spatial perspectives. As one of the first studies to examine pupils' spatial thinking within the educational use of data physicalization, this investigation illustrated how spatial thinking and hands-on making evolved together through the embodied learning process.

An open-ended, design-by-analogy project was then developed to examine how such tasks engage pupils in spatial thinking processes (Chapter 3). Pupils abstracted visual-spatial features from biological inspirations, inferred form-function relationships, and creatively transformed these features into original designs. Through the use of mental and visual representations, they explored design possibilities, envisioned scenarios to assess functionality, and identified opportunities for iteration and refinement. This study further highlighted potential differences between spatial thinking processes involved in solving open-ended design tasks and in closed-ended spatial tests. Whereas spatial ability tests typically require the analysis and mental transformation of predefined features of test items, creative design problem-solving necessitates broader retrieval, exploration, and evaluation of diverse visual-spatial information. Although the role of spatial ability in analytical STEM problems is increasingly understood, its role in creative, open-ended problem-solving remains less explored. Clarifying this relationship is essential for developing spatial training that supports the creation of innovative and functional solutions to real-world challenges.

Insights from these two case studies, together with emerging literature, underscore the meaningful role that spatial thinking plays in design processes. To extend this line of inquiry, a mixed-methods study on the evaluation of creativity in design ideas and prototypes (Chapter 4) and a quantitative analysis of how spatial ability contributes to creative design outcomes (Chapter 5) were conducted. In Chapter 4, using the method of comparative judgement, expert judges holistically evaluated the creativity of a range of design ideas and prototypes developed by 201 pupils aged 10 to 14. Judges frequently emphasized novelty while also considering usability, presentation, idea qualities, feasibility, and problem-solving seen in the ideas and prototypes. The coding scheme derived from judges' comments aligned closely with established frameworks for assessing product creativity, supporting the validity of comparative judgement as a holistic assessment approach and offering nuanced insights into complex constructs such as creativity and design quality.

Using comparative judgement results as proxies for pupils' design ideation and design prototype performance, Chapter 5 further examined whether and how spatial ability relates to pupils' creative design performance—an area that has received limited attention in prior research. The analyses also explored whether the contribution of spatial ability was unique or overlapped with creativity measures, including divergent thinking and creative imagery. Despite small effect sizes, spatial ability correlated significantly with both design ideation and design prototype quality. After controlling for age, verbal ability, and creativity, spatial ability accounted for unique variance in design prototype quality but not in ideation quality. This suggested that spatial ability plays a more prominent role during prototyping than during ideation, potentially by supporting the development and refinement of design solutions that are both novel and useful. In addition, spatial ability showed a small but significant correlation with divergent thinking, pointing to potential shared cognitive underpinnings that warrant further investigation. In contrast, no significant relationship was found between spatial ability and creative imagery, reinforcing the conceptual distinction between spatial and object-based imagery. Taken together, these findings offered new evidence linking domain-general spatial ability with creative design performance among school-age pupils and contribute to the ongoing discussion about the relationship between spatial ability and creativity.

6.3 General discussion

6.3.1 Merging qualitative and quantitative evidence: potential mechanisms linking spatial thinking to design

Design involves multiple stages, and this dissertation primarily examined design ideation and prototyping. The following discussion integrates qualitative and quantitative evidence to outline potential mechanisms through which spatial thinking may support pupils' design performance, as well as areas that warrant further investigation.

Spatial thinking and problem representation

Design begins with framing and reframing the design problem, with novelty emerging through the interplay between evolving problem definitions and tentative solutions (Dorst & Cross, 2001). Although it might be expected that verbal ability plays a central role in understanding design briefs, analyses in Chapter 5 revealed no significant association between pupils' verbal ability and either ideation or prototype quality. This may be partly due to the fact that the design task itself was easy to comprehend, such that advanced vocabulary offered no meaningful advantage. Spatial ability, by contrast, not only correlated significantly with both creative design ideation and prototype quality but also uniquely predicted design prototype scores after controlling for creativity measures. This suggests that spatial ability may have played a role in supporting the translation of words into mental representations that could be central to design problem-solving.

Prior research lends support to this interpretation. Antonietti (1999) showed that when tackling practical problems involving interpreting problem descriptions and imagining the manipulation or arrangement of objects, mental representations play a crucial role in helping individuals visualize the problem and identify the relationships among different elements in problem descriptions. Moreover, Antonietti proposed that those who are aware of the usefulness of mental representations tend to make more effective use of this tool in problem-solving. A more recent study by Duffy et al. (2020) demonstrated that spatial ability supports the structuring of information from verbal mathematical descriptions into mental representations linked to the knowledge schemas needed for solving open-ended mathematical problems. Duffy et al. (2024) further found that mental representations mediated the relationship between undergraduate students' spatial ability and their performance on mathematics word problems. Building on these findings, together with the results from Chapter 5, this dissertation proposes that pupils with stronger spatial skills may be better able to recognize when a design brief should be "re-pictured." That is, they may be more effective at converting verbal descriptions into visual-spatial mental representations of design problems and linking these representations to design-specific knowledge schemas, such as using multiple planes of view to communicate a design or illustrating the intended function of a design in comic-strip style. These creative uses of design schemas and representational practices appear to enrich pupils' design outcomes in terms of both visual and technical details.

Future research is needed to test this hypothesis by examining whether problem representation ability mediates the relationship between spatial ability and design performance. If supported, this would highlight design education as a valuable context for investigating how mental representations support the activation and application of domain-specific knowledge schemas.

Spatial thinking, perspective taking, and social understanding

A second potential mechanism concerns the relationship between spatial ability and social-cognitive skills. There is growing evidence that stronger spatial skills are associated with greater social competence (Erle & Topolinski, 2015; Tanaś & Szarek, 2021; Tian et al., 2021; Tsomokos & Flouri, 2025). Erle and Topolinski (2015) reported a positive correlation between adults' visuospatial perspective-taking and their ability to empathically adopt others' perspectives, suggesting that imagining how someone feels may involve mentally simulating their spatial viewpoints. Among children, Tanaś and Szarek (2021) found that 5- to 6-year-olds' skills in mental transformation, mental folding, and visuospatial perspective-taking were all significantly associated with social understanding skills, including empathy and psychological perspective-taking. Furthermore, visuospatial perspective-taking skills significantly predicted both social understanding and prosocial behaviors.

These findings provide an additional lens for interpreting the association between spatial ability and design performance. In design education, students are explicitly encouraged to consider users' needs and experiences, which involves empathetic perspective taking (Klapwijk & Van Doorn, 2015). For instance, biomedical design challenges involving users with disabilities require designers to mentally or physically simulate everyday activities from the users' viewpoints (Klapwijk, 2017). Similarly, the SnackSafe on the Beach design task (Chapter 4, Figure 4.1) required pupils to envision the users' experience in an environment surrounded by seagulls and to devise solutions that balanced effectiveness, user-friendliness, and animal welfare. Such tasks draw on both empathetic and spatial perspective-taking.

If future research continues to support the link between spatial and social perspective-taking, this relationship may help explain how spatial thinking contributes to design performance. Meanwhile, design education—particularly its increasing focus on human-centered design (IDEO.org, 2015; Klapwijk & Van Doorn, 2015)—presents valuable opportunities to cultivate spatial skills that are closely intertwined with social understanding.

Spatial thinking as a resource for novel and functional design

A third potential mechanism linking spatial ability to design performance concerns pupils' capacity to integrate novelty with functional viability during prototyping. Analyses of judges' comments in Chapter 4 revealed that although originality was consistently valued, considerations related to usability, functionality, and user experiences were mentioned even more frequently when evaluating design prototypes. Importantly, judges did not always separate creativity from usefulness; instead, they often emphasized how a design's affordances contributed to its perceived creativity. This pattern reflects a broader consensus in design research that creative products must balance novelty with effectiveness (Cropley & Cropley, 2010).

Considerations of functionality may not always appear as salient visuo-spatial features of a design; rather, designers need to associate visual information with assumptions about use, environment, and interaction in order to evaluate whether a design will function effectively in practice (Suwa et al., 1998). Evidence of this process among younger learners can be seen in pupils' design work showcased in Chapter 3. For example, when developing a camping tarp design, two pupils revised their initial sketch into an elevated pyramid structure after visualizing how rainfall would interact with the surface in possible usage scenarios. Such episodes illustrate how spatial thinking supports pupils in linking form, function, and context during iterative design development.

The relationship between spatial ability and functional aspects of design is further supported by quantitative evidence. Previous studies have reported associations between spatial ability and functionality-related dimensions of design performance (Chang, 2014; Raju et al., 2023). Together with the findings of this dissertation, this body of work underscores spatial ability's role in supporting the functional refinement of creative ideas.

By enabling designers to anticipate interactions among components, materials, users, and environments, and to form mental representations of how artefacts will be experienced, operated, and situated in real-world contexts, spatial thinking may support higher-order design thinking that coordinates form, function, and contextual constraints into coherent solutions. Future research should further examine how spatial thinking is engaged during prototyping, testing, and iterative refinement, particularly in relation to functional evaluation.

6.3.2 Complexities in measuring spatial and design performance[6]

Quantitative examinations of the relationship between spatial ability and design performance remain inconclusive. Whereas some studies have reported no significant association between spatial ability and design outcomes (Allen, 2010; Cho, 2017; Reid & Sorby, 2023), others have found that only domain-specific spatial tests—rather than domain-general spatial measures—correlated with architecture or interior design performance (Cho & Suh, 2019; Suh & Cho, 2020). By contrast, Chang (2014) and the findings reported in Chapter 5 of this dissertation showed that domain-general spatial tests correlated significantly with design performance, with small-to-medium effect sizes.

Measurement challenges in assessing design and spatial performance

These discrepancies may be partly attributed to the complexities involved in measuring these constructs. To measure design performance, studies have employed different approaches: some focused primarily on end products (Chang, 2014; Suh & Cho, 2020), others have evaluated both processes and products (Cho, 2017; Kimbell & Stables, 2007), and some have relied on course grades as indicators of design performance (Sutton & Williams, 2010). Additional concerns have been raised regarding the validity of the design tasks used to measure design performance. For instance, Allen (2010) noted that the design task used in their study may not have adequately assessed design creativity due to unclear success criteria.

Similar challenges arise in the measurement of spatial ability. There have been long-standing debates about whether spatial tests necessarily measure spatial ability. Researchers have argued that performance on spatial tests is often influenced by other cognitive (Toth & Campbell, 2019), social (Bartlett & Camba, 2023), or performance-related factors, such as familiarity with using test-taking strategies (Hegarty, 2018). Moreover, the types of stimuli used in spatial tests may introduce biases in comprehension and response processes (Caissie et al., 2009; Fisher et al., 2018). Others have pointed out that existing spatial ability tests do not sufficiently address all the spatial skills needed for solving spatially complex STEM problems (Atit et al., 2020; Uttal et al., 2024). For instance, penetrative thinking is a critical skill for geology students to visualize spatial relations inside objects and to understand microstructures in minerals (Ormand et al., 2014). This domain-specific skill—related to but distinct from commonly assessed spatial

[6] *This section includes text that has previously appeared, with modifications, in the following source:*

Zhu, C., & Klapwijk, R. (2024, June). The Spatial Aspect of Designing: Opportunities, Challenges, and Conjectures on Engaging Pupils in Spatial Thinking Through Design Education. In *German Conference on Spatial Cognition* (pp. 97-113). Cham: Springer Nature Switzerland.

skills such as visualization or mental rotation—is rarely reflected in standard spatial tests. Similarly, Suh and Cho (2020) found that students' design performance correlated only with spatial tasks specifically developed for interior design, but not with general spatial measures. More recently, Bartlett and Camba (2024) highlighted that current spatial tests tend to cover a narrow set of skills and fail to address the spatial skills needed to visualize 3D garment construction in apparel design. Taken together, these limitations in the scope and specificity of spatial measures may help explain the inconsistent findings regarding the relationship between spatial ability and design performance in the literature.

Task structure and cognitive demands in spatial and design tasks

Another perspective for interpreting these discrepancies concerns the nature of the problems assessed by conventional spatial tests. Given that spatial ability involves "the ability to make use of simulated mental imagery to solve problems'' (Schneider & McGrew, 2012, p. 129), it is essential to consider the kinds of problems being addressed, whether they are open-ended or closed-ended. Many widely used multiple-choice spatial tests primarily engage cognitive and problem-solving processes needed for solving closed-ended problems. Empirical evidence suggests that even in mathematics education—a field whose relationship with spatial ability is relatively well understood—the association between spatial ability and open-ended mathematics problem solving varies by age. For example, spatial ability did not significantly predict third graders' performance on open-ended mathematics problems (Bahar & Maker, 2015), whereas it significantly predicted eighth graders' performance, particularly for challenging open-ended math problems (Wang et al., 2022).

Designers' ways of problem-solving are arguably different from those in natural or social sciences (Cross, 2006; Lawson, 1979). This difference primarily stems from design problems being ill-defined and open-ended, necessitating both divergent and convergent thinking (Goldschmidt, 2016; Schut et al., 2020). Specifically, designers need divergent thinking to 'shift perspectives on existing information (seeing it in a new way) or transforming it,' thereby producing 'alternative or multiple solutions' and developing products that differ from those that existed before (Cropley & Cropley, 2009, p. 67). In contrast, conventional intelligence and achievement tests emphasize accuracy and convergence on a single correct solution.

It is therefore reasonable to conjecture that designers employ problem-solving strategies that have nuanced differences from those used to solve analytical problems, and that certain aspects of spatial thinking may be unique to the design contexts, and are neither assessed nor developed in conventional spatial tasks or training. Through an in-depth qualitative analysis of the types of spatial reasoning exhibited by fifth- and sixth-graders during a series of STEAM challenges, Ramey and colleagues (2020) found that many forms of spatial reasoning employed in authentic design and making contexts were not adequately reflected in conventional psychometric assessments of spatial ability. Building on these insights and the findings of this dissertation, we propose two additional distinctions between psychometric spatial test problems and open-ended design problems, which may inform future investigation of pupils' spatial ability within design and STEM education contexts.

From selection to generation and iteration: key differences between spatial tests and design tasks

The first distinction lies in the fact that closed-ended tasks typically require choosing from predefined options, whereas open-ended tasks involve generating multiple plausible solutions before narrowing down to one. The assessment of creativity provides a useful parallel: it cannot be measured through selecting predetermined answers. Well-known creativity measures, such as the Alternate Uses Task (Guilford, 1967) and the Torrance Tests of Creative Thinking (Torrance, 1974), emphasize the generation of original and varied solutions. Similarly, design problems rarely offer ready-made solutions. Instead, they require divergent idea generation before a potentially fitting solution can be developed.

Contrarily, many spatial ability assessments predominantly feature well-defined and closed-ended questions. In widely used diagnostic spatial tests, including mental rotation tasks (Vandenberg & Kuse, 1978) and the Purdue Spatial Visualization Test (Guay, 1977), and more recent measures such as the Spatial Reasoning Instrument (Ramful et al., 2017) and the Mental Folding Test for Children (Harris et al., 2013), test-takers are presented with a limited set of options and are expected to identify only one or a few correct answers. Designers, however, bring what has not yet been made or seen before into reality (Wilson & Harris, 2004), meaning that they engage in imagining and creating a multitude of options in their minds (e.g., Figure 3.4 in Chapter 3). Envisioning and developing a range of tentative mental representations of ideas differs from choosing among visible options on paper or screen.

Additionally, there are several known strategies for solving typical spatial tasks presented in multiple-choice forms. For instance, one might concentrate on envisioning the transformation of a certain part of the stimuli, rather than the entire stimuli, to reduce the complexity of mental transformation required (Gluck & Fitting, 2003). One may also rely on test-taking strategies such as ruling out incorrect answers one by one (Hegarty, 2018) or selectively verifying likely answers (Maresch, 2014). Although these strategies can be effective in tasks with certain response formats, their usefulness is limited in open-ended contexts.

A second distinction between closed-ended and open-ended tasks concerns the emphasis on arriving at a correct solution versus seeking iterative improvements. In typical spatial tests, once a predetermined solution is identified, there is often no need for further improvement. However, iteration is a central feature of design practice and is strongly encouraged in design education (Adams & Atman, 1999; Kimbell & Stables, 2007; Looijenga et al., 2015). After generating multiple ideas, designers do not simply settle on a desirable option. Rather, they continually refine and further elaborate on their solutions. This iterative process is fundamental to producing designs that are both functional and innovative.

Implications for spatial ability research and design education

In summary, while designers use fundamental spatial skills such as mental rotation and spatial visualization to imagine possible design configurations and improvements, the ways in which these skills are employed—and the goals they serve—might differ from their use in completing spatial tests. Current assessments of spatial ability and existing approaches to spatial training may not fully account for these nuanced differences, and the construct of spatial thinking does not yet adequately reflect the creative and iterative use of spatial skills in design contexts. Hence, outcomes from existing spatial ability tests may not adequately capture individuals' ability to use spatial skills in addressing open-ended, real-world design problems.

Learning from how designers use spatial thinking could therefore offer valuable insights for spatial ability research by advancing understanding of how spatial thinking is applied in diverse and authentic contexts, particularly in envisioning and shaping the technologies and environments that surround us. A more comprehensive understanding of how spatial thinking develops through open-ended design activities may inform the design of instructional approaches that better prepare pupils to tackle complex STEM problems. Future studies are needed to further investigate how spatial ability supports open-ended problem-solving across disciplines and to develop assessments that align more closely with classroom realities and educational needs.

6.4 Limitations and future directions

A primary limitation of this dissertation is that it did not systematically examine the influence of individual-level variables, such as gender and socioeconomic status (SES), on pupils' spatial ability. These factors may have influenced how pupils respond to spatial tests, spatial interventions, or spatially enriched teaching practices.

There remains a gender gap in STEM participation between male and female students (Stoet & Geary, 2018). Gender differences in spatial ability have long been debated with regard to their causes (Bartlett & Camba, 2023; Linn & Petersen, 1985; Lippa et al., 2010) and their implications for STEM performance (Sorby, 2009; Wang & Degol, 2017). Recent evidence suggests that these differences may depend on the nature of assessment tasks. Bartlett and Camba (2024) demonstrated that while male students outperformed female students in a classic spatial ability test that has been shown to reliably predict STEM performance, female students outperformed male students in an apparel design spatial ability test specifically assessing the types of visualization and transformation needed for 3D garment construction. This finding supports the view that observed gender differences may be partly due to the type and format of stimuli used in spatial tests. Additional factors, such as time limit (Maeda & Yoon, 2013), test-taking strategies (Hegarty, 2018), or stereotype threat (Neuburger et al., 2015), may also influence performance on spatial tests and warrant further investigation. At present, limited quantitative research has examined the influence of gender on the role of spatial ability in design performance (e.g., Cho, 2017; Reid & Sorby, 2023). Future studies are needed to further clarify how gender may moderate the relationship between spatial ability and design performance, and how these patterns compare with those observed in other STEM disciplines. Future research should also be mindful of selecting spatial tests and assessment procedures that measure students' spatial ability without perpetuating the gender gap.

Greater attention is also needed to explore how creative design and making practices that have traditionally been gendered might support spatial skills development. For example, fiber arts have been identified as potentially beneficial for fostering spatial skills needed to visualize transformations within and between objects (Bennett-Pierre & Gunderson, 2023). Adopting more inclusive approaches that leverage diverse creative practices may help bridge gender-related gaps in spatial ability training and contribute to more equitable outcomes in STEM education.

A second limitation concerns the sampling context. All studies in this dissertation were conducted in international schools in the Netherlands. The voluntary participation of schools and classes may have introduced self-selection bias, as participating teachers may have been particularly interested in design education or confident in their pupils' abilities. Besides, pupils attending international schools may not adequately represent the socioeconomic diversity of the broader student population. Beyond gender disparities, the underrepresentation of students from low socioeconomic backgrounds, ethnic minorities, and other marginalized groups in STEM education remains a persistent concern (MacPhee et al., 2013; Saw et al., 2018). Research has consistently shown that children from low-SES backgrounds perform worse on spatial tasks than their middle- and high-SES peers (Casey et al., 2011; Jirout & Newcombe, 2015; Levine et al., 2005), with such differences emerging as early as preschool (Bower et al., 2020). These disparities are often attributed to unequal access to spatially enriching toys, play activities, and learning opportunities in everyday environments (Casey et al., 2011).

Importantly, spatial skills are malleable, and pupils from low-SES backgrounds can benefit substantially from targeted spatial training (Bower et al., 2020; Bower et al., 2022). Although recruiting diverse and representative samples can be challenging due to demographic and practical constraints, recent studies demonstrate increasing efforts to address this issue (Jirout & Newcombe, 2015; Bower et al., 2025). Future research should therefore prioritize more representative sampling strategies and explicitly examine the moderating roles of gender and socioeconomic status when designing and evaluating spatial interventions. By accounting for these individual and contextual factors, future work can enhance the generalizability of findings and deepen the understanding of equity issues in spatial skills development.

6.4.1 Recommendations for future research and educational practice

An exploration of spatial thinking in design goes beyond theoretical inquiries and invites us to reimagine educational practices by intertwining spatial thinking with creative and interdisciplinary problem-solving. Across both qualitative and quantitative examinations, it became evident that a wide range of spatial skills are frequently engaged in design-based practices. These findings highlight the potential to understand and foster spatial skills not only through analytical STEM problems but also through open-ended, design-based problems, thereby offering an alternative perspective on how spatial thinking can support authentic STEM learning.

Scaffolding spatial thinking through design tools and resources

An increasing number of studies have explored how digital technologies support spatial thinking, including computer-aided design (Fowler et al., 2024), 3D printing (Bhaduri et al., 2021), and virtual and augmented reality (González, 2018). However, such technologies are not always accessible or suitable across design stages and contexts.

Currently, there are various non-digital resources available to support design processes, including the YourTurn toolbox developed at TU Delft (Klapwijk et al., 2021), the design portfolio and formative assessment tools from the Technology Education Research Unit (Kimbell & Stables, 2007), and the Field Guide to Human-Centered Design and the Design Thinking for Educators Toolkit from IDEO.org (2015; n.d.). Yet, these tools do not currently make the use of spatial skills explicit.

Future research could explore how design toolkits can be developed to more deliberately scaffold spatial skills development in design. One possible direction is to focus on metacognitive awareness of the importance of spatial skills in designing. Prior work suggests that learners who recognize the value of mental representations are better able to use them effectively in problem solving (Antonietti, 1999). Instructional supports for spatial thinking could therefore be woven into various stages of the design process—

for example, by guiding learners to create visual overviews of users' needs, use visual diagrams or models to map elements of a design problem, and build prototypes that provide immediate visual or physical feedback.

Leveraging open-ended design problems for spatial and STEM learning

Educational research has long emphasized the value of authentic learning experiences that engage learners in ill-defined, open-ended, and real-world problems (Herrington et al., 2014; Herrington & Oliver, 2000). Within STEM education, growing attention has been directed toward fostering creativity, divergent thinking, and problem solving through open-ended tasks, including in mathematics (Kwon et al., 2006), physics (Kolodner et al., 2003), life science (Hmelo et al., 2014), and chemistry (de Lavoir et al., 2023; Stammes et al., 2020). For instance, sixth graders' iterative development of artificial lung designs facilitated their understanding of the respiratory system (Hmelo et al., 2014), while tenth graders' prototypes on long-lasting soap bubbles supported structure-property reasoning and micro-macro thinking in chemistry (de Lavoir et al., 2023). In each case, learners engaged in spatial thinking when visualizing systems, structures, and relationships.

Similarly, research on spatial ability has emphasized the importance of developing spatial skills to address real-world problems (Atit et al., 2020; Liben et al., 2002; Uttal & Cohen, 2012). The National Research Council (2006) described spatial thinking as 'using the properties of space as a vehicle for structuring problems, for finding answers, and for expressing solutions' (p. 3). Because design-based thinking and practices are frequently used to generate innovative and contextually appropriate solutions in science, technology, and engineering (Kimbell & Stables, 2007), and because design activities support the development of spatial skills and the understanding of spatial relations (Carbonell-Carrera et al., 2020; Dere & Kalelioglu, 2020; Lin, 2016), design-based practices may serve as a bridge between spatial interventions and authentic STEM learning experiences.

Future research may consider exploring (1) the potential of incorporating open-ended design problems into spatial interventions, (2) how participation in design activities supports the development and application of specific spatial thinking strategies, and (3) educational approaches for equipping learners with the spatial skills needed to represent, test, and iterate on ideas in two- and three-dimensional media, both by hand and using digital technologies, ultimately producing novel and functional solutions to real-world design and STEM problems.

Design as a bridge for interdisciplinary knowledge and cognitive skills

Design is typically characterized by its interdisciplinary nature, drawing on knowledge and skills from multiple domains (Kimbell & Perry, 2001). Empirical research indicates that design-based activities promote the development of transferable skills across disciplines. For instance, Levin and Verner (2020) demonstrated that 3D design and printing workshops supported seventh-graders' analytical thinking and applied mathematical thinking, while Bicer et al. (2017) reported that computer-aided design activities helped high school students develop more positive perceptions of creativity and problem-solving skills in STEM. Together, these examples illustrate that the versatile nature of design makes it a unique context for cultivating not only spatial skills but also domain-specific and domain-general knowledge and competencies for STEM learning.

Beyond spatial skills, the design process recruits a range of other cognitive functions, such as working memory capacity (Bilda & Gero, 2007) and analogical reasoning (Hey et al., 2008) during idea generation, perspective taking to infer users' needs (Klapwijk & Van Doorn, 2015), and convergent thinking for evaluating concepts (Goldschmidt, 2016). Future research could investigate how designers utilize spatial skills alongside other cognitive skills to formulate design solutions, and the extent to which these skills contribute to the quality of design products, the effectiveness of design processes, and designers' overall competencies.

Practical implications for design educators

Design educators can draw on the findings of this dissertation to identify where spatial thinking likely takes place in pupils' design processes, and how it may be challenged or scaffolded through instructional practices. In addition to prior research suggesting the meaningful role spatial ability plays in modeling, iterating, and testing the design (Bhaduri et al., 2021; Fowler et al., 2024; Ramey & Uttal, 2017; Ramey et al., 2020), the findings reported in Chapter 5 indicate that spatial ability is also important in design ideation and early prototyping. Building on these insights, this dissertation offers practical suggestions for scaffolding pupils' spatial thinking throughout the design cycle, including problem framing, idea generation, prototype building, testing, and iteration.

To stimulate visual-spatial reasoning during problem framing and idea generation, teachers may select design challenges that can be approached from different angles rather than tasks with a limited number of known solutions. Such ill-defined problems encourage pupils to explore and visualize in their minds a broader range of conceptual and form possibilities. Providing a rich and varied set of visual examples related to the design problem and guiding their attention toward underlying form-function relationships (Casakin & Goldschmidt, 2000) may further help pupils extract transferable ideas. Additionally, embedding perspective-taking exercises can encourage pupils to shift from their own perspective toward envisioning user experiences in the given context, while deepening their understanding of human-centered design.

Since designing embodies constant interactions between the hand and the mind (Kimbell & Stables, 2007), it is essential to provide materials, tools, and sufficient time throughout the design cycle for pupils to externalize their ideas through sketches, drawings, and tangible or virtual artifacts. During prototyping, testing, and iteration, teachers can leverage a multimodal approach by combining hands-on making with digital fabrication (Pedgley & Şener, 2025). The requirement to transition between dimensions, such as translating a 2D sketch into a 3D clay model, or a 3D computer-aided design into a 2D technical drawing, actively challenges pupils' capacity for spatial transformation and spatial scaling, while providing direct visual feedback that supports iterative reflection. Rather than merely evaluating whether a prototype functions, teachers can prompt pupils to examine the appearance and functionality of their prototypes from different spatial points of view. Even when prototypes do not function as intended, pupils can develop valuable spatial insights by reconciling the discrepancies between their artifacts and their mental representations (Khunyakari et al., 2007; Ramey et al., 2020; Zhu, Klapwijk et al., 2023).

Ultimately, incorporating spatial language, gestures, sketches, analogies, as well as both hands-on and virtual modeling into everyday instruction provides pupils with cognitive tools for forming spatial representations and reasoning about spatial information (Newcombe, 2017). By integrating these practices across the design process, educators can better support pupils in translating abstract ideas into novel, viable, and spatially coherent solutions.

References

Acer, D. (2016). Children and architecture: architectural design education for young children in Turkey. Childhood Education, 92(2), 99-110.

Adams, R. S., & Atman, C. J. (1999). Cognitive processes in iterative design behavior. In FIE'99 Frontiers in Education. 29th Annual Frontiers in Education Conference. Designing the Future of Science and Engineering Education. Conference Proceedings (IEEE Cat. No. 99CH37011 (Vol. 1, pp. 11A6-13). IEEE.

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. Learning and instruction, 16(3), 183-198.

Akayuure, P., Asiedu-Addo, S. K., & Alebna, V. (2016). Investigating the effect of origami instruction on preservice teachers' spatial ability and geometric knowledge for teaching. International Journal of Education in Mathematics, Science and Technology, 4(3), 198–209.

Allen, A. D. (2010). Complex spatial skills: The link between visualization and creativity. Creativity Research Journal, 22(3), 241-249.

Amabile, T. M., & Gitomer, J. (1984). Children's artistic creativity: Effects of choice in task materials. Personality and Social Psychology Bulletin, 10(2), 209-215.

Antonietti, A. (1999). Can students predict when imagery will allow them to discover the problem solution?. European Journal of Cognitive Psychology, 11(3), 407-428.

Antonietti, A., & Colombo, B. (2011). Mental imagery as a strategy to enhance creativity in children. Imagination, Cognition and Personality, 31(1), 63-77.

Asghar, M., Sorby, S. A., Reid, C., & Raju, G. (2024). Eye-Tracking Analysis of Problem-Solving Behavior in Design Tasks in Undergraduate Engineering: A Comparison of High and Low Spatial Visualizers. In 2024 ASEE Annual Conference & Exposition.

Athavankar, U. A. (1997). Mental imagery as a design tool. Cybernetics & Systems, 28(1), 25-42.

Athavankar, U. A. (1999). Gestures, imagery and spatial reasoning. In J. S. Gero & B. Tversky (Eds.), Visual and spatial reasoning in design (pp. 103–128). Key Centre of Design Computing and Cognition, University of Sydney.

Atit, K., Power, J. R., Pigott, T., Lee, J., Geer, E. A., Uttal, D. H., Ganley, C. M. & Sorby, S. A. (2022). Examining the relations between spatial skills and mathematical performance: A meta-analysis. Psychonomic bulletin & review 29, 699–720.

Atit, K., Uttal, D. H., & Stieff, M. (2020). Situating space: Using a discipline-focused lens to examine spatial thinking skills. Cognitive Research: Principles and Implications, 5(1), 19.

Atkinson, S. (2000). Does the need for high levels of performance curtail the development of creativity in design and technology project work?. International Journal of Technology and Design Education, 10, 255-281.

Aziz-Zadeh, L., Liew, S. L., & Dandekar, F. (2013). Exploring the neural correlates of visual creativity. Social Cognitive and Affective Neuroscience, 8(4), 475–480.

Bae, S. S., Vanukuru, R., Yang, R., Gyory, P., Zhou, R., Do, E. Y. L., & Szafir, D. A. (2022). Cultivating visualization literacy for children through curiosity and play. IEEE Transactions on Visualization and Computer Graphics, 29(1), 257-267.

Bahar, A., & Maker, C. J. (2015). Cognitive backgrounds of problem solving: A comparison of open-ended vs. closed mathematics problems. Eurasia Journal of Mathematics, Science and Technology Education, 11(6), 1531-1546.

Baker, J., Jones, D., & Burkman, J. (2009). Using visual representations of data to enhance

sensemaking in data exploration tasks. Journal of the Association for Information Systems, 10(7), 2.
Bakker, A. (2018). Design research in education: A practical guide for early career researchers (1st ed.). Routledge.
Ball, L. J., & Christensen, B. T. (2019). Advancing an understanding of design cognition and design metacognition: Progress and prospects. Design Studies, 65, 35-59.
Barratt, E. S. (1953). An analysis of verbal reports of solving spatial problems as an aid in defining spatial factors. The Journal of Psychology, 36(1), 17–25.
Bartholomew, S. R., Mentzer, N., Jones, M., Sherman, D., & Baniya, S. (2022). Learning by evaluating (LbE) through adaptive comparative judgment. International Journal of Technology and Design Education, 32(2), 1191-1205
Bartholomew, S. R., Nadelson, L. S., Goodridge, W. H., & Reeve, E. M. (2018). Adaptive comparative judgment as a tool for assessing open-ended design problems and model eliciting activities. Educational Assessment, 23(2), 85-101.
Bartholomew, S. R., Ruesch, E. Y., Hartell, E., & Strimel, G. J. (2020). Identifying design values across countries through adaptive comparative judgment. International Journal of Technology and Design Education, 30(2), 321-347.
Bartholomew, S. R., Strimel, G. J., & Yoshikawa, E. (2019). Using adaptive comparative judgment for student formative feedback and learning during a middle school design project. International Journal of Technology and Design Education, 29, 363-385.
Bartholomew, S., & Yoshikawa-Ruesch, E. (2018). A systematic review of research around adaptive comparative judgement (ACJ) in K-16 education. CTETE-Research Monograph Series, 1(1), 6-28.
Bartlett, K. A., & Camba, J. D. (2023). Gender differences in spatial ability: A critical review. Educational Psychology Review, 35(1), 8.
Bartlett, K. A., & Camba, J. D. (2024). Toward a broader understanding of spatial ability in CAD education. Computer-Aided Design and Applications, 24(1), 39-54.
Barsalou, L. W. (2008). Grounded cognition. Annual Review Psychology, 59(1), 617-645.
Batic, J. (2011). The field trip as part of spatial (architectural) design art classes. CEPS journal, 1(2), 73-86.
Baykal, G. E., Alaca, I. V., Yantaç, A. E., & Göksun, T. (2018). A review on complementary natures of tangible user interfaces (TUIs) and early spatial learning. International journal of child-computer interaction, 16, 104-113.
Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. Memory & cognition, 42, 1186-1197.
Becker, J. P., & Shimada, S. (1997). The Open-Ended Approach: A New Proposal for Teaching Mathemat- ics. National Council of Teachers of Mathematics.
Bejar, I. I. (2012). Rater cognition: Implications for validity. Educational Measurement: Issues and Practice, 31(3), 2-9.
Ben-Chaim, D., Lappan, G., & Houang, R. T. (1988). The effect of instruction on spatial visualization skills of middle school boys and girls. American Educational Research Journal, 25(1), 51-71.
Benedek, M., Jauk, E., Fink, A., Koschutnig, K., Reishofer, G., Ebner, F., & Neubauer, A. C. (2014). To cre- ate or to recall? Neural mechanisms underlying the generation of creative new ideas. NeuroImage, 88, 125–133.
Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. Intelligence, 46, 73-

83.
Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological), 57(1), 289-300.
BenMahmoud-Jouini, S., & Midler, C. (2020). Unpacking the notion of prototype archetypes in the early phase of an innovation process. Creativity and Innovation Management, 29(1), 49-71.
Bennett-Pierre, G., & Gunderson, E. A. (2023). Fiber arts require spatial skills: How a stereotypically feminine practice can help us understand spatial skills and improve spatial learning. Sex Roles, 88(1), 1-16.
Benson, C., & Lunt, J. (2011). We're creative on a Friday afternoon: Investigating children's perceptions of their experience of design & technology in relation to creativity. Journal of Science Education and Technology, 20, 679-687.
Benton, T. (2021). Comparative Judgement for Linking Two Existing Scales. Frontiers in Education, 6.
Benyus, J. M. (1997). Biomimicry: Innovation inspired by nature (p. 320). New York: Morrow.
Ben-Zvi, D., & Garfield, J. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. The challenge of developing statistical literacy, reasoning and thinking, 66, 3-15.
Bergey, B. W., Cromley, J. G., & Newcombe, N. S. (2015). Teaching high school biology students to coordinate text and diagrams: relations with transfer, effort, and spatial skill. International Journal of Science Education, 37(15), 2476-2502.
Beveridge, M., & Parkins, E. (1987). Visual representation in analogical problem solving. Memory & Cognition, 15, 230–237.
Bhaduri, S., Biddy, Q. L., Bush, J., Suresh, A., & Sumner, T. (2021). 3DnST: A framework towards understanding children's interaction with Tinkercad and enhancing spatial thinking skills. In Proceedings of the 20th annual ACM interaction design and children conference (pp. 257-267).
Bhaduri, S., Van Horne, K., & Sumner, T. (2019). Designing an informal learning curriculum to develop 3D modeling knowledge and improve spatial thinking skills. In Extended abstracts of the 2019 CHI conference on human factors in computing systems (pp. 1-8).
Bhargava, R., & D'Ignazio, C. (2017). Data sculptures as a playful and low-tech introduction to working with data. In Proceedings of the 2017 Designing Interactive Systems. Association for Computing Machinery.
Bicer, A., Nite, S. B., Capraro, R. M., Barroso, L. R., Capraro, M. M., & Lee, Y. (2017). Moving from STEM to STEAM: The effects of informal STEM learning on students' creativity and problem solving skills with 3D printing. In 2017 IEEE frontiers in education conference (FIE) (pp. 1-6). IEEE.
Bilda, Z., & Gero, J. S. (2007). The impact of working memory limitations on the design process during conceptualization. Design Studies, 28(4), 343-367.
Biomimicry 3.8. (2015). Biomimicry DesignLens: A visual guide. Accessed on January 18, 2022, from https://www.biomimicry.net/the-buzz/resources/biomimicry-designlens/
Blikstein, P. (2018). Maker movement in education: History and prospects. Handbook of technology educa- tion (pp. 419–437). Springer.
Bockmon, R., Cooper, S., Gratch, J., Zhang, J., & Dorodchi, M. (2020). Can students' spatial skills predict their programming abilities?. In Proceedings of the 2020 ACM

Conference on Innovation and Technology in Computer Science Education (pp. 446-451).
Boden, M. (2001). Creativity and knowledge. Creativity in education (pp. 95–102). Continuum.
Borst, G., & Kosslyn, S. M. (2010). Individual differences in spatial mental imagery. Quarterly Journal of Experimental Psychology, 63(10), 2031-2050.
Bower, C. A., Zimmermann, L., Verdine, B. N., Pritulsky, C., Golinkoff, R. M., & Hirsh-Pasek, K. (2022). Enhancing spatial skills of preschoolers from under-resourced backgrounds: A comparison of digital app vs. Concrete materials. Developmental Science, 25(1), e13148.
Bower, C. A., Zimmermann, L., Verdine, B. N., Toub, T.S., Hirsh-Pasek, K., & Golinkoff, R.M. (2025). What does play have to do with it? A concrete and digital spatial intervention with 3-year-olds predicts spatial and math learning. Developmental Psychology.
Bower, C., Zimmermann, L., Verdine, B., Toub, T. S., Islam, S., Foster, L., Evans, N., Odean, R., Cibischino, A., Pritulsky, C., Hirsh-Pasek, K. & Golinkoff, R. M. (2020). Piecing together the role of a spatial assembly intervention in preschoolers' spatial and mathematics learning: Influences of gesture, spatial language, and socioeconomic status. Developmental Psychology, 56(4), 686.
Boyatzis, R. (1998). Transforming qualitative information: Thematic analysis and code development. Sage.
Bramley, T., & Vitello, S. (2019). The effect of adaptivity on the reliability coefficient in adaptive comparative judgement. Assessment in Education: Principles, Policy & Practice, 26(1), 43-58.
Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101.
Bryan, L. A., Moore, T. J., Johnson, C. C., & Roehrig, G. H. (2015). Integrated STEM education. In STEM road map (pp. 23-38). Routledge.
Buckley, J. (2024a). An Open-Source Adaptive Comparative Judgement App for Technology Education Research and Practice: Alpha Version. Journal of Technology Education, 36(1), 58-82.
Buckley, J. (2024b). Adaptive Comparative Judgement App. Accessed on January 15, 2025, from https://jeffbuckley1992.shinyapps.io/comparative_judgement/
Buckley, J. (2024c). Conducting Power Analyses to Determine Sample Sizes in Quantitative Research: A Primer for Technology Education Researchers Using Common Statistical Tests. Journal of Technology Education, 35(2), 81-109.
Buckley, J., & Canty, D. (2022). Assessing performance: Addressing the technical challenge of comparing novel portfolios to the "ACJ-Steady State." PATT39: PATT on the Edge - Technology, Innovation and Education, 523–537.
Buckley, J., Canty, D., & Seery, N. (2022). An exploration into the criteria used in assessing design activities with adaptive comparative judgment in technology education. Irish Educational Studies, 41(2), 313-331.
Buckley, J., Seery, N., & Canty, D. (2018). A heuristic framework of spatial ability: A review and synthesis of spatial factor literature to support its translation into STEM education. Educational Psychology Review, 30(3), 947-972.
Buckley, J., Seery, N., Canty, D., & Gumaelius, L. (2022). The importance of spatial ability within technology education. Applications of research in technology education: Helping teachers develop research- informed practice (pp. 165–182). Springer Nature Singapore.

Buckley, J., Seery, N., & Kimbell, R. (2023). Modelling approaches to combining and comparing independent adaptive comparative judgement ranks. In The 40th International Pupils’ Attitudes Towards Technology Conference Proceedings 2023 (Vol. 1, No. October).
Burghardt, M. D., Hecht, D., Russo, M., Lauckhardt, J., & Hacker, M. (2010). A study of mathematics infusion in middle school technology education classes. Journal of Technology Education, 22(1), 58–74.
Burte, H., Gardony, A. L., Hutton, A., & Taylor, H. A. (2017). Think3d!: Improving mathematics learning through embodied spatial training. Cognitive Research: Principles and Implications, 2(1), 13.
Burton, L. J., & Fogarty, G. J. (2003). The factor structure of visual imagery and spatial abilities. Intelligence, 31(3), 289-318.
Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. Journal of Educational Psychology, 79(4), 474.
Caissie, A. F., Vigneau, F., & Bors, D. A. (2009). What does the mental rotation test measure? An Analysis of Item Difficulty and Item Characteristics. The Open Psychology Journal, 2(1), 94–102.
Caprile, M., Palmén, R., Sanz, P., & Dente, G. (2015). Encouraging STEM studies labour market situation and comparison of practices targeted at young people in different member states. Policy Department A, 12.
Carbonell-Carrera, C., Saorin, J. L., Melián-Díaz, D., & Hess-Medler, S. (2020). Spatial orientation skill performance with a workshop based on green infrastructure in cities. ISPRS International Journal of Geo-Information, 9(4), 216.
Card, S. K., & Mackinlay, J. (1997). The structure of the information visualization design space. In Proceedings of VIZ'97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium (pp. 92-99). IEEE.
Carroll, J. B. (1993). Human cognitive abilities: A survey of factor-analytical studies. New York: Cambridge University Press.
Casakin, H. (2004). Visual analogy as a cognitive strategy in the design process. Expert versus novice performance. Journal of Design Research, 4(2), 197–217.
Casakin, H. (2008). Factors of design problem-solving and their contribution to creativity. Open house international, 33(1), 46-60.
Casakin, H., Davidovitch, N., & Milgram, R. M. (2010). Creative thinking as a predictor of creative problem solving in architectural design students. Psychology of Aesthetics, Creativity, and the Arts, 4(1), 31–35.
Casakin, H., & Goldschmidt, G. (1999). Expertise and the use of visual analogy: Implications for design education. Design Studies, 20(2), 153–175.
Casakin, H. P., & Goldschmidt, G. (2000). Reasoning by visual analogy in design problem-solving: the role of guidance. Environment and Planning B: Planning and Design, 27(1), 105-119.
Casakin, H., & Kreitler, S. (2005). The nature of creativity in design. In J. S. Gero and N. Bonnardel (Eds.), Studying designers, 5, (pp. 87-100). Key Centre of Design Computing and Cognition, University of Sydney.
Casakin, H., & Wodehouse, A. (2021). A systematic review of design creativity in the architectural design studio. Buildings, 11(1), 31.
Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., & Copley, J. (2008). The development of spatial skills through interventions involving block building activities. Cognition and instruction, 26(3), 269-309.

Casey, B. M., Dearing, E., Vasilyeva, M., Ganley, C. M., & Tine, M. (2011). Spatial and numerical predictors of measurement performance: The moderating effects of community income and gender. Journal of educational psychology, 103(2), 296.

Casey, B. M., Pezaris, E., Fineman, B., Pollock, A., Demers, L., & Dearing, E. (2015). A longitudinal analysis of early spatial skills compared to arithmetic and verbal skills as predictors of fifth-grade girls' math reasoning. Learning and Individual Differences, 40, 90-100.

Chang, Y. S. (2014). 3D-CAD effects on creative design performance of different spatial abilities students. Journal of Computer Assisted Learning, 30(5), 397-407.

Chao, S. J., Stigler, J. W., & Woodward, J. A. (2000). The effects of physical materials on kindergartners' learning of number concepts. Cognition and Instruction, 18(3), 285-316.

Cheng, Y. L., & Mix, K. S. (2014). Spatial training improves children’s mathematics ability. Journal of Cognition and Development, 15(1), 2–11.

Charyton, C., Jagacinski, R. J., Merrill, J. A., Clifton, W., & DeDios, S. (2011). Assessing creativity specific to engineering with the revised creative engineering design assessment. Journal of Engineering Education, 100(4), 778-799.

Cho, J. Y. (2017). An investigation of design studio performance in relation to creativity, spatial ability, and visual cognitive style. Thinking Skills and Creativity, 23, 67-78.

Cho, J. Y., & Suh, J. (2019). Understanding spatial ability in interior design education: 2D–to–3D visualization proficiency as a predictor of design performance. Journal of Interior Design, 44(3), 141-159.

Christensen, B. T., & Ball, L. J. (2016). Creative analogy use in a heterogeneous design team: The pervasive role of background domain knowledge. Design Studies, 46, 38-58.

Christensen, B. T., & Schunn, C. D. (2009). The role and impact of mental simulation in design. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 23(3), 327-344.

Christiaans, H. H. (2002). Creativity as a design criterion. Communication Research Journal, 14(1), 41-54.

Chrysikou, E. G. (2014). Creative states: A cognitive neuroscience approach to understanding and improving creativity in design. Studying visual and spatial reasoning for design creativity (pp. 227–243). Springer.

Chrysikou, E. G., & Thompson-Schill, S. L. (2011). Dissociable brain states linked to common and creative object use. Human brain mapping, 32(4), 665-675.

Chrysikou, E. G., & Weisberg, R. W. (2005). Following the wrong footsteps: fixation effects of pictorial examples in a design problem-solving task. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(5), 1134.

Claxton, A. F., Pannells, T. C., & Rhoads, P. A. (2005). Developmental trends in the creativity of school-age children. Creativity Research Journal, 17(4), 327-335.

Clement, J. (1981). Analogy generation in scientific problem solving. In Proceedings of the Third Annual Conference of the Cognitive Science Society.

Clement, J. (1986). Methods for evaluating the validity of hypothesized analogies. In Proceedings of the Eighth Annual Meeting of the Cognitive Science Society. Lawrence Erlbaum Associates.

Clement, J. (1988). Observed methods for generating analogies in scientific problem solving. Cognitive Science, 12(4), 563–586.

Clement, J. J. (Ed.). (2008). Creative model construction in scientists and students. Springer Netherlands.

Clement, J. (2009). Analogical reasoning via imagery: the role of transformations and simulations. New frontiers in analogy research (pp. 463–472). New Bulgarian University Press.
Cohrssen, C., de Quadros-Wander, B., Page, J., & Klarin, S. (2017). Between the big trees: A project-based approach to investigating shape and spatial thinking in a kindergarten program. Australasian Journal of Early Childhood, 42(1), 94-104.
Cohrssen, C., & Pearn, C. (2021). Assessing preschool children's maps against the first four levels of the primary curriculum: Lessons to learn. Mathematics Education Research Journal, 33(1), 43-60.
Cook, K. L., & Bush, S. B. (2018). Design thinking in integrated STEAM learning: Surveying the landscape and exploring exemplars in elementary grades. School Science and Mathematics, 118(3-4), 93-103.
Cooper, M. M., Stieff, M., & DeSutter, D. (2017). Sketching the invisible to predict the visible: From drawing to modeling in chemistry. Topics in cognitive science, 9(4), 902-920.
Coxon, S. V. (2012). Innovative allies: Spatial and creative abilities. Gifted Child Today, 35(4), 277-284.
Craft, A. (2001). Little c creativity. In A. Craft, B. Jeffrey, & M. Leibling (Eds.), Creativity in education, 45–61. London: Continuum.
Cremin, T., Glauert, E., Craft, A., Compton, A., & Stylianidou, F. (2018). Creative little scientists: Exploring pedagogical synergies between inquiry-based and creative approaches in early years science. In Creativity and creative pedagogies in the early and primary years (pp. 45-60). Routledge.
Crismond, D. P., & Adams, R. S. (2012). The informed design teaching & learning matrix. Journal of Engineering Education-Washington, 101(4), 738.
Cromley, J. G., Bergey, B. W., Fitzhugh, S., Newcombe, N., Wills, T. W., Shipley, T. F., & Tanaka, J. C. (2013). Effects of three diagram instruction methods on transfer of diagram comprehension skills: The critical role of inference while learning. Learning and Instruction, 26, 45-58.
Cropley, A. J. (2005). Creativity and problem-solving: Implications for classroom assessment. Leicester: British Psychological Society.
Cropley, A. (2006). In praise of convergent thinking. Creativity Research Journal, 18(3), 391–404.
Cropley, A. J., & Cropley, D. (2009). Fostering creativity: A diagnostic approach for higher education and organizations. Cresskill, NJ: Hampton Press.
Cropley, D., & Cropley, A. (2010). Recognizing and fostering creativity in technological design education. International Journal of Technology and Design Education, 20, 345-358.
Cropley, D. H., & Kaufman, J. C. (2012). Measuring functional creativity: Non-expert raters and the Creative Solution Diagnosis Scale. The journal of creative behavior, 46(2), 119-137.
Cross, N. (2006). Designerly ways of knowing (pp. 1-13). Springer London.
Custer, R. L. (1995). Examining the dimensions of technology. International journal of technology and design education, 5, 219-244.
Dasgupta, S., & Hill, B. M. (2017, May). Scratch community blocks: Supporting children as data scientists. In Proceedings of the 2017 CHI conference on human factors in computing systems (pp. 3620-3631).
Daugherty, J., & Mentzer, N. (2008). Analogical reasoning in the engineering design process and technology education applications. Journal of Technology Education,

19(2), 7.
Davies, S. R., & Horst, M. (2016). Science communication: Culture, identity and citizenship. Springer.
de la Harpe, B., Peterson, J. F., Frankham, N., Zehner, R., Neale, D., Musgrave, E., & McDermott, R. (2009). Assessment focus in studio: What is most prominent in architecture, art and design?. International Journal of Art & Design Education, 28(1), 37-51.
de Lavoir, S. V. P., Otter, M. J. D., de Vries, M., & Barendsen, E. (2023). Characterising Structure-Property Reasoning within a Chemical Design Challenge:'Green Bubble Soap'. In The 40th International Pupils' Attitudes Towards Technology Conference Proceedings.
de Vries, M. J. (2021). Design-based learning in science and technology as integrated STEM. In Design-based concept learning in science and technology education (pp. 14-24). Brill.
Department for Education. (2013). National curriculum in England: design and technology programmes of study. Accessed on April 3, 2025, from https://www.gov.uk/government/publications/national-curriculum-in-england-design-and-technology-programmes-of-study
Dere, H. E., & Kalelioglu, F. (2020). The effects of using web-based 3D design environment on spatial visualisation and mental rotation abilities of secondary school students. Informatics in Education, 19(3), 399-424.
Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. Educational Researcher, 32(1), 5–8.
Diezmann, C. M., & Watters, J. J. (2000). Identifying and supporting spatial intelligence in young children. Contemporary Issues in Early Childhood, 1(3), 299-313.
Dilling, F., & Vogler, A. (2021). Fostering spatial ability through computer-aided design: A case study. Digital Experiences in Mathematics Education, 7(2), 323-336.
Dorst, K., & Cross, N. (2001). Creativity in the design process: co-evolution of problem–solution. Design studies, 22(5), 425-437.
Dortch, D., & Patel, C. (2017). Black undergraduate women and their sense of belonging in STEM at predominantly White institutions. NASPA Journal About Women in Higher Education, 10(2), 202-215.
D'souza, N. S. (2010). The metaphor of an ensemble: design creativity as skill integration. In Design Creativity 2010 (pp. 281-288). London: Springer London.
Duffy, G., Sorby, S., & Bowe, B. (2020). An investigation of the role of spatial ability in representing and solving word problems among engineering students. Journal of Engineering Education, 109(3), 424-442.
Duffy, G., Sorby, S., & Bowe, B. (2024, June). Exploring the role of spatial ability in the mental representation of word problems in mathematics. In Frontiers in Education (Vol. 9, p. 1346474). Frontiers Media SA.
Dunn, L. M., & Dunn, D. M. (2007). PPVT-4: Peabody picture vocabulary test. Bloomington, MN: Pearson Assessments.
Dziedziewicz, D., & Karwowski, M. (2015). Development of children's creative visual imagination: A theoretical model and enhancement programmes. Education 3-13, 43:4, 382-392.
Egelandsdal, K., Hartell, E., & Færstad, J. O. (2025). Exploring the practical feasibility of adaptive comparative judgment as a summative assessment method. Assessment & Evaluation in Higher Education, 1–16.
Ehrlich, S. B., Levine, S. C., & Goldin-Meadow, S. (2006). The importance of gesture in

children's spatial reasoning. Developmental Psychology, 42(6), 1259.
Ekstrom, R.B., French, J.W., Harman, H.H., Dermen, D. (1976). Manual for kit of factor-referenced cognitive tests. Educational Testing Service, Princeton, NJ.
English, L. D. (2019). Learning while designing in a fourth-grade integrated STEM problem. International Journal of Technology and Design Education, 29(5), 1011-1032.
Erle, T. M., & Topolinski, S. (2015). Spatial and empathic perspective-taking correlate on a dispositional level. Social Cognition, 33(3), 187-210.
Fan, S. C., & Yu, K. C. (2017). How an integrative STEM curriculum can benefit students in engineering design practices. International Journal of Technology and Design Education, 27, 107-129.
Fatourou, P., Papageorgiou, Y., & Petousi, V. (2019). Women are needed in STEM: European policies and incentives. Communications of the ACM, 62(4), 52-52.
Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. Health services research, 48(6 Pt 2), 2134-2156.
Ferguson, E. S. (1977). The Mind's Eye: Nonverbal Thought in Technology: " Thinking with pictures" is an essential strand in the intellectual history of technological development. Science, 197(4306), 827-836.
Finke, R. A. (1989). Principles of mental imagery. The MIT Press.
Finke, R. A. (1990). Creative imagery: Discoveries and inventions in visualization. Hillsdale, NJ: Erlbaum.
Finke, R. A., Pinker, S., & Farah, M. J. (1989). Reinterpreting visual patterns in mental imagery. Cognitive Science, 13, 51–78.
Finke, R. A., Ward, T. B., & Smith, S. M. (1992). Creative cognition: Theory, research, and applications. The MIT Press.
Fisher, M. L., Meredith, T., & Gray, M. (2018). Sex differences in mental rotation ability are a consequence of procedure and artificiality of stimuli. Evolutionary Psychological Science, 4, 124-133.
Fish, J., & Scrivener, S. (1990). Amplifying the mind's eye: sketching and visual cognition. Leonardo, 23(1), 117-126.
Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educational and psychological measurement, 33(3), 613-619.
Forthmann, B., Gerwig, A., Holling, H., Çelik, P., Storme, M., & Lubart, T. (2016). The be-creative effect in divergent thinking: The interplay of instruction and object frequency. Intelligence, 57, 25-32.
Forthmann, B., Holling, H., Zandi, N., Gerwig, A., Çelik, P., Storme, M., & Lubart, T. (2017). Missing creativity: The effect of cognitive workload on rater (dis-) agreement in subjective divergent-thinking scores. Thinking Skills and Creativity, 23, 129-139.
Forthmann, B., Jendryczko, D., Scharfen, J., Kleinkorres, R., Benedek, M., & Holling, H. (2019). Creative ideation, broad retrieval ability, and processing speed: A confirmatory study of nested cognitive abilities. Intelligence, 75, 59–72.
Fowler, S., Kennedy, J. P., Cutting, C., Gabriel, F., & Leonard, S. N. (2024). Self-determined learning in a virtual makerspace: A pathway to improving spatial reasoning for upper primary students. International Journal of Technology and Design Education, 34(2), 563-584.
Francis, K., Rothschuh, S., Poscente, D., & Davis, B. (2022). Malleability of spatial reasoning with short-term and long-term robotics interventions. Technology,

Knowledge and Learning, 27(3), 927-956.
Frick, A., Daum, M. M., Wilson, M., & Wilkening, F. (2009). Effects of action on children's and adults' mental imagery. Journal of Experimental Child Psychology, 104(1), 34–51.
Frick, A., & Newcombe, N. S. (2015). Young children's perception of diagrammatic representations. Spatial Cognition & Computation, 15(4), 227-245.
Fu, K., Murphy, J., Yang, M., Otto, K., Jensen, D., & Wood, K. (2015). Design-by-analogy: Experimental evaluation of a functional analogy search methodology for concept generation improvement. Research in Engineering Design, 26, 77–95.
Gabora, L. (2010). Revenge of the "neurds": Characterizing creative thought in terms of the structure and dynamics of memory. Creativity Research Journal, 22(1), 1–13.
Gagnier, K. M., Atit, K., Ormand, C. J., & Shipley, T. F. (2017). Comprehending 3D diagrams: Sketching to support spatial reasoning. Topics in Cognitive Science, 9(4), 883–901.
Gencer, A. S., Doğan, H., & Bilen, K. (2020). Developing biomimicry STEM activity by querying the relationship between structure and function in organisms. Turkish Journal of Education, 9(1), 64–105.
Gentner, D. (1987). Mechanisms of analogical learning (Vol. 87, No. 1381). Department of Computer Science, University of Illinois at Urbana-Champaign.
Gentner, D., & Medina, J. (1998). Similarity and the development of rules. Cognition, 65(2–3), 263–297.
Gentner, D., Levine, S. C., Ping, R., Isaia, A., Dhillon, S., Bradley, C., & Honke, G. (2016). Rapid learning in a children's museum via analogical comparison. Cognitive Science, 40(1), 224–240.
Gero, J. S. (1990). Design prototypes: a knowledge representation schema for design. AI magazine, 11(4), 26-26.
Gero, J. S. (1996). Creativity, emergence and evolution in design. Knowledge-Based Systems, 9(7), 435–448.
Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. Cognitive Psychology, 12(3), 306–355.
Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. Cognitive psychology, 15(1), 1–38.
Gilligan, K. A., Hodgkiss, A., Thomas, M. S., & Farran, E. K. (2019). The developmental relations between spatial cognition and mathematics in primary school children. Developmental Science, 22(4), e12786.
Gilligan, K. A., Thomas, M. S., & Farran, E. K. (2020). First demonstration of effective spatial training for near transfer to spatial performance and far transfer to a range of mathematics skills at 8 years. Developmental Science, 23(4), e12909.
Gluck, J., & Fitting, S. (2003). Spatial strategy selection: Interesting incremental information. International Journal of Testing, 3(3), 293-308.
Goel, A. K. (1997). Design, analogy, and creativity. IEEE Expert, 12(3), 62–70.
Goktepe Yildiz, S., & Ozdemir, A. S. (2020). The effects of engineering design processes on spatial abilities of middle school students. International Journal of Technology and Design Education, 30, 127-148.
Goldschmidt, G. (1994). On visual design thinking: The vis kids of architecture. Design Studies, 15(2), 158–174.
Goldschmidt, G. (1995). Visual displays for design: Imagery, analogy and databases of visual images. Visual Databases in Architecture, 1, 53–74.
Goldschmidt, G. (2001). Visual analogy—a strategy for design reasoning and learning.

Design knowing and learning: Cognition in design education (pp. 199–219). Elsevier Science.
Goldschmidt, G. (2016). Linkographic evidence for concurrent divergent and convergent thinking in creative design. Creativity research journal, 28(2), 115-122.
Goldschmidt, G., & Tatsa, D. (2005). How good are good ideas: Correlates of design creativity. Design Studies, 26, 593–611.
Goldman, S., Carroll, M., & Royalty, A. (2009). Destination, imagination & the fires within: Design thinking in a middle school classroom. In Proceedings of the seventh ACM conference on Creativity and cognition (pp. 371-372).
González, N. A. A. (2018). Development of spatial skills with virtual reality and augmented reality. International Journal on Interactive Design and Manufacturing (IJIDeM), 12, 133-144.
Gough, A. (2015). STEM policy and science education: Scientistic curriculum and sociopolitical silences. Cultural Studies of Science Education, 10, 445-458.
Green, A. E., Kraemer, D. J., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2012). Neural correlates of creativity in analogical reasoning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 38(2), 264.
Groth, C., & Mäkelä, M. (2016). The knowing body in material exploration. Studies in Material Thinking, 14(2), 1-11.
Guay, R. (1977). Purdue Spatial Visualization Test: Visualization of Rotations. Purdue Research Foundation.
Guignard, J. H., & Lubart, T. I. (2007). A comparative study of convergent and divergent thinking in intellectually gifted children. Gifted and Talented International, 22(1), 9-15.
Guilford, J. P. (1967). The nature of human intelligence, New York: McGraw-Hill.
Guilford, J. P. (1968). Intelligence, creativity, and their educational implications. Robert R. Knapp.
Guo, S., Wang, X., Deng, W., Hong, J., Wang, J., & Wu, Y. (2022). Whose spatial ability benefits from learning with 3D design? From the perspective of learning analysis. Educational Technology & Society, 25(1), 179-192.
Hallström, J., & Ankiewicz, P. (2023). Design as the basis for integrated STEM education: A philosophical framework. In Frontiers in Education (Vol. 8, p. 1078313). Frontiers Media SA.
Hambrick, D. Z., Libarkin, J. C., Petcovic, H. L., Baker, K. M., Elkins, J., Callahan, C. N., Turner, S. P., Rench, T. A., & LaDue, N. D. (2012). A test of the circumvention-of-limits hypothesis in scientific problem solving: The case of geological bedrock mapping. Journal of Experimental Psychology: General, 141(3), 397–403.
Hannafin, R. D., Truxaw, M. P., Vermillion, J. R., & Liu, Y. (2008). Effects of spatial ability and instructional program on geometry achievement. The Journal of Educational Research, 101(3), 148-157.
Harle, M., & Towns, M. (2011). A review of spatial ability literature, its connection to chemistry, and implications for instruction. Journal of Chemical Education, 88(3), 351-360.
Harris, J., Newcombe, N. S., & Hirsh-Pasek, K. (2013). A new twist on studying the development of dynamic spatial transformations: Mental paper folding in young children. Mind, Brain, and Education, 7(1), 49-55.
Hartell, E., & Buckley, J. (2021). Comparative judgment: An overview. Handbook for Online Learning Contexts: Digital, Mobile and Open: Policy and Practice, 289-307.
Hawes, Z., & Ansari, D. (2020). What explains the relationship between spatial and

mathematical skills? A review of evidence from brain and behavior. Psychonomic bulletin & review, 27, 465-482.
Hawes, Z. C., Gilligan-Lee, K. A., & Mix, K. S. (2022). Effects of spatial training on mathematics performance: A meta-analysis. Developmental Psychology, 58(1), 112.
Hawes, Z., Moss, J., Caswell, B., Naqvi, S., & MacKinnon, S. (2017). Enhancing children's spatial and numerical skills through a dynamic spatial approach to early geometry instruction: Effects of a 32-week intervention. Cognition and Instruction, 35(3), 236–264.
Hawes, Z., Moss, J., Caswell, B., & Poliszczuk, D. (2015). Effects of mental rotation training on children's spatial and mathematics performance: A randomized controlled study. Trends in Neuroscience and Education, 4(3), 60–68.
Hegarty, M. (2010). Components of spatial intelligence. In Psychology of learning and motivation (Vol. 52, pp. 265-297). Academic Press.
Hegarty, M. (2014). Spatial thinking in undergraduate science education. Spatial Cognition & Computation, 14(2), 142-167.
Hegarty, M. (2018). Ability and sex differences in spatial thinking: What does the mental rotation test really measure?. Psychonomic bulletin & review, 25, 1212-1219.
Hegarty, M., Keehner, M., Cohen, C., Montello, D. R., & Lippa, Y. (2007). The role of spatial cognition in medicine: Applications for selecting and training professionals. Applied spatial cognition: From research to cognitive technology (pp. 285–315). Lawrence Erlbaum Associates Publishers.
Hegarty, M., & Sims, V. K. (1994). Individual differences in mental animation during mechanical reasoning. Memory & Cognition, 22, 411-430.
Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. Educational technology research and development, 48(3), 23-48.
Herrington, J., Reeves, T. C., & Oliver, R. (2014). Authentic learning environments. Handbook of research on educational communications and technology, 401-412.
Hey, J., Linsey, J., Agogino, A. M., & Wood, K. L. (2008). Analogies and metaphors in creative design. International Journal of Engineering Education, 24(2), 283.
Hinze, S. R., Williamson, V. M., Shultz, M. J., Deslongchamps, G., Williamson, K. C., & Rapp, D. N. (2014). Spatial ability and learning from visualizations in STEM disciplines. Space in mind: Concepts for spatial learning and education, 99-118.
Hmelo, C. E., Holton, D. L., & Kolodner, J. L. (2014). Designing to learn about complex systems. In Design education (pp. 247-298). Routledge.
Hodgkiss, A., Gilligan, K. A., Tolmie, A. K., Thomas, M. S., & Farran, E. K. (2018). Spatial cognition and science achievement: The contribution of intrinsic and extrinsic spatial skills from 7 to 11 years. British Journal of Educational Psychology, 88(4), 675-697.
Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—a meta-analytic review. Educational psychology review, 22(3), 245-269.
Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics, 6(2), 65–70.
Hoskens, M., & Wilson, M. (2001). Real-time feedback on rater drift in constructed-response items: An example from the Golden State Examination. Journal of Educational Measurement, 38(2), 121-145.
Houde, S., & Hill, C. (1997). What do prototypes prototype?. In Handbook of human-computer interaction (pp. 367-381). North-Holland.
Howard, T. J., Culley, S. J., & Dekoninck, E. (2008). Describing the creative design

process by the integration of engineering design and cognitive psychology literature. Design studies, 29(2), 160-180.
Hsi, S., Linn, M. C., & Bell, J. E. (1997). The role of spatial reasoning in engineering and the design of spatial instruction. Journal of Engineering Education, 86(2), 151–158.
Huk, T. (2006). Who benefits from learning with 3D models? The case of spatial ability. Journal of computer assisted learning, 22(6), 392-404.
Hull, C., & Willett, W. (2017, May). Building with data: Architectural models as inspiration for data physicalization. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (pp. 1217-1264).
Huron, S., Carpendale, S., Boy, J., & Fekete, J. D. (2016). Using viskit: A manual for running a constructive visualization workshop. In Pedagogy of Data Visualization Workshop at IEEE VIS 2016.
Huron, S., Carpendale, S., Thudt, A., Tang, A., & Mauerer, M. (2014). Constructive visualization. In Proceedings of the 2014 conference on Designing interactive systems (pp. 433-442).
Huron, S., Gourlet, P., Hinrichs, U., Hogan, T., & Jansen, Y. (2017). Let's get physical: Promoting data physicalization in workshop formats. In Proceedings of the 2017 Conference on Designing Interactive Systems (pp. 1409-1422).
Huron, S., Jansen, Y., & Carpendale, S. (2014). Constructing visual representations: Investigating the use of tangible tokens. IEEE transactions on visualization and computer graphics, 20(12), 2102-2111.
Hurtienne, J., & Reinhardt, D. (2017). Teaching data physicalisation to HCI students. In Proc. pedagogy Physicalization workshop DIS.
Huttenlocher, J. (1968). Constructing spatial images: A strategy in reasoning. Psychological Review, 75(6), 550–560.
IDEO.org. (2015). The Field Guide to Human-Centered Design. Accessed on July 10, 2023, from https://www.designkit.org/resources/1.html
IDEO.org. (n.d.). Design thinking for educators toolkit. Accessed on May 20, 2024, from https://page.ideo.com/design-thinking-edu-toolkit
Inhelder, B., & Piaget, J. (1964). The early growth of logic in the child. New York: Harper & Row.
International Baccalaureate. (2009). Mathematics scope and sequence. International Baccalaureate Organization. Cardiff, Wales.
International Baccalaureate Organization. (n.d.). Middle Years Programme. https://www.ibo.org/myp/
Isham, D. (1997). Developing a computerized interactive visualization assessment. Journal of Computer- Aided Environmental Design and Education, 3(1), 1–15.
Jaison, D., Watson, K., Merzdorf, H. E., Williford, B., White, L. L. A., Douglas, K. A., & Hammond, T. A. (2021). Assessing drawing self-efficacy: A validation study using exploratory factor analysis (efa) for the drawing self-efficacy instrument (dsei). In 2021 ASEE Virtual Annual Conference Content Access.
Jankowska, D. M., Gajda, A., & Karwowski, M. (2019). How children's creative visual imagination and creative thinking relate to their representation of space. International Journal of Science Education, 41(8), 1096-1117.
Jankowska, D. M., & Karwowski, M. (2015). Measuring creative imagery abilities. Frontiers in Psychology, 6, 159933.
Jankowska, D. M., & Karwowski, M. (2020). Test of creative imagery abilities. Wydawnictwo Liberi Libri.
Jansen, Y., & Dragicevic, P. (2013). An interaction model for visualizations beyond the

desktop. IEEE Transactions on Visualization and Computer Graphics, 19(12), 2396-2405.
Jansen, Y., Dragicevic, P., & Fekete, J. D. (2013). Evaluating the efficiency of physical visualizations. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 2593-2602).
Jansen, Y., Dragicevic, P., Isenberg, P., Alexander, J., Karnik, A., Kildal, J., Subramanian, S., & Hornbæk, K. (2015). Opportunities and challenges for data physicalization. In proceedings of the 33rd annual acm conference on human factors in computing systems (pp. 3227-3236).
Jansen, Y., & Hornbæk, K. (2015). A psychophysical investigation of size as a physical variable. IEEE transactions on visualization and computer graphics, 22(1), 479-488.
Jauk, E., Benedek, M., & Neubauer, A. C. (2014). The road to creative achievement: A latent variable model of ability and personality predictors. European journal of personality, 28(1), 95-105.
Jirout, J. J., & Newcombe, N. S. (2015). Building blocks for developing spatial skills: Evidence from a large, representative US sample. Psychological science, 26(3), 302-310.
Jones, I., & Alcock, L. (2014). Peer assessment without assessment criteria. Studies in Higher Education, 39(10), 1774-1787.
Jones, I., & Inglis, M. (2015). The problem of assessing problem solving: Can comparative judgement help?. Educational Studies in Mathematics, 89, 337-355.
Julià, C., & Antolì, J. Ò. (2018). Enhancing spatial ability and mechanical reasoning through a STEM course. International Journal of Technology and Design Education, 28(4), 957–983.
Kanis, M. (2019). Physical sensemaking: Crafting for an invisible world of data. In Workshop: Troubling Innovation: Craft and Computing Across Boundaries [the weekend before] The ACM CHI Conference on Human Factors in Computing Systems.
Kavakli, M., & Gero, J. S. (2001). Sketching as mental imagery processing. Design Studies, 22(4), 347–364.
Kavakli, M., & Gero, J. S. (2002). The structure of concurrent cognitive actions: A case study on novice and expert designers. Design Studies, 23(1), 25–40
Kell, H. J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2013). Creativity and technical innovation spatial ability's unique role. Psychological Science, 24(9), 1831–1836.
Kelly, A. V., Kimbell, R. A., Patterson, V. J., Saxton, J., & Stables, K. (1987). Design and technological activity: a framework for assessment. London: HM Stationery Office, Department of Education and Science, Assessment of Performance Unit.
Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. International Journal of STEM education, 3, 1-11.
Kelley, T. R., Sung, E., Han, J., & Knowles, J. G. (2023). Impacting secondary students' STEM knowledge through collaborative STEM teacher partnerships. International Journal of Technology and Design Education, 33(4), 1563-1584.
Khunyakari, R., Mehrotra, S., Chunawala, S., & Natarajan, C. (2007). Design and technology productions among middle school students: an Indian experience. International Journal of Technology and Design Education, 17(1), 5-22.
Kim, K. H. (2008). Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores. The Journal of Creative Behavior, 42(2), 106-130.
Kim, M. J., & Maher, M. L. (2008). The impact of tangible user interfaces on spatial cognition during collaborative design. Design Studies, 29(3), 222-253.

Kimbell, R. (2012). Evolving project e-scape for national assessment. International Journal of Technology and Design Education, 22, 135-155.
Kimbell, R., & Perry, D. (2001). Design and technology in a knowledge economy. London: Engineering Council.
Kimbell, R., & Stables, K. (2007). Researching design learning: Issues and findings from two decades of research and development. Springer Dordrecht.
Kimbell, R., Stables, K., & Green, R. (1996). Understanding practice in design and technology. Open University Press.
Kirsh, D. (2009). Interaction, external representation and sense making. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 31, No. 31).
Klapwijk, R. (2017). Creativity in design. In Teaching design and technology creatively (pp. 51-72). Routledge.
Klapwijk, R., Gielen, M., Schut, A., van Mechelen, M., & Stables, K. (2021). Your Turn for the teacher: Guidebook to develop real-life design lessons for use with 8-14-year-old pupils. Access: https://studiolab.ide.tudelft.nl/studiolab/codesignwithkids/files /Your-Turn-for-the-teacher-Guidebook.pdf
Klapwijk, R. M., & Stables, K. (2023). Design Learning: Pedagogic Strategies That Enable Learners to Develop Their Design Capability. In Bloomsbury handbook of technology education (pp. 271-289). Bloomsbury.
Klapwijk, R., & Van Doorn, F. (2015). Contextmapping in primary design and technology education: a fruitful method to develop empathy for and insight in user needs. International Journal of Technology and Design Education, 25(2), 151-167.
Kleibeuker, S. W., Koolschijn, P. C. M., Jolles, D. D., Schel, M. A., De Dreu, C. K., & Crone, E. A. (2013). Prefrontal cortex involvement in creative problem solving in middle adolescence and adulthood. Developmental cognitive neuroscience, 5, 197-206.
Koh, J. H. L., Chai, C. S., Wong, B. & Hong, H. Y. (2015). Design thinking and education (pp. 1-15). Springer Singapore.
Kokotovich, V., & Purcell, T. (2000). Mental synthesis and creativity in design: an experimental examination. Design studies, 21(5), 437-449.
Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntambekar, S. & Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design (tm) into practice. The journal of the learning sciences, 12(4), 495-547.
Kontra, C., Lyons, D. J., Fischer, S. M., & Beilock, S. L. (2015). Physical experience enhances science learning. Psychological science, 26(6), 737-749.
Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of chiropractic medicine, 15(2), 155-163.
Kosslyn, S. M. (1978). The representational-development hypothesis. In P. A. Ornstein (Ed.), Memory development in children (pp. 157–189). Lawrence Erlbaum.
Kosslyn, S. M. (1980). Image and mind. Harvard University Press.
Kosslyn, S. M., Pinker, S., Smith, G. E., & Shwartz, S. P. (1979). On the demystification of mental imagery. Behavioral and Brain Sciences, 2(4), 535–548.
Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. Memory & cognition, 33(4), 710-726.
Kozhevnikov, M., Kozhevnikov, M., Yu, C. J., & Blazhenkova, O. (2013). Creativity, visualization abilities, and visual cognitive style. British journal of educational psychology, 83(2), 196-209.

Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2015). Students' ability to solve process-diagram problems in secondary biology education. Journal of Biological Education, 49(1), 91-103.
Kudrowitz, B. M., & Wallace, D. (2013). Assessing the quality of ideas from prolific, early-stage product ideation. Journal of Engineering Design, 24(2), 120-139.
Kus, M., & Cakiroglu, E. (2022). Mathematics in the informal setting of an art studio: students' visuospatial thinking processes in a studio thinking-based environment. Educational Studies in Mathematics, 110(3), 545-571.
Kwon, O. N., Park, J. H., & Park, J. S. (2006). Cultivating divergent thinking in mathematics through an open-ended approach. Asia Pacific Education Review, 7, 51-61.
Lamon, S. J. (1996). The development of unitizing: Its role in children's partitioning strategies. Journal for Research in Mathematics Education, 27(2), 170-193.
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. biometrics, 159-174.
Lane, D., & Sorby, S. (2022). Bridging the gap: Blending spatial skills instruction into a technology teacher preparation programme. International Journal of Technology and Design Education, 32(4), 2195–2215.
Lawson, B. (2005). How designers think: The design process demystified (4th ed.). Routledge.
Lawson, B. R. (1979). Cognitive strategies in architectural design. Ergonomics, 22(1), 59-68.
Leopold, C., & Leutner, D. (2012). Science text comprehension: Drawing, main idea selection, and summarizing as learning strategies. Learning and Instruction, 22(1), 16-26.
Lesterhuis, M., Bouwer, R., Van Daal, T., Donche, V., & De Maeyer, S. (2022). Validity of comparative judgment scores: How assessors evaluate aspects of text quality when comparing argumentative texts. In Frontiers in Education (Vol. 7, p. 823895). Frontiers Media SA.
Lesterhuis, M., Verhavert, S., Coertjens, L., Donche, V., & De Maeyer, S. (2017). Comparative judgement as a promising alternative to score competences. In Innovative practices for higher education assessment and measurement (pp. 119-138). IGI Global.
Levin, L., & Verner, I. M. (2020). Fostering students' analytical thinking and applied mathematical skills through 3D design and printing. In 2020 IEEE global engineering education conference (EDUCON) (pp. 145-149). IEEE.
Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. Psychological science, 16(11), 841-845.
Lewis, T. (2006). Creativity: A framework for the design/problem solving discourse in technology education, Journal of Technology Education, 17(1), 36-53.
Lewis, T. (2009). Creativity in technology education: Providing children with glimpses of their inventive potential. International Journal of Technology and Design Education, 19, 255-268.
Liebschner, J. (2006). A child's work: freedom and guidance in Froebel's educational theory and practice.
Liben, L. S., Kastens, K. A., & Stevenson, L. M. (2002). Real-world knowledge through real-world maps: A developmental guide for navigating the educational terrain. Developmental Review, 22(2), 267-322.

Lin, H. (2016). Influence of design training and spatial solution strategies on spatial ability performance. International Journal of Technology and Design Education, 26, 123-131.
Lin, X. (1992). Visualization for the document space. In Proceedings Visualization'92 (pp. 274-281). IEEE.
Lindström, L. (2006). Creativity: What is it? Can you assess it? Can it be taught?. International Journal of Art & Design Education, 25(1), 53-66.
Link, T., Moeller, K., Huber, S., Fischer, U., & Nuerk, H. C. (2013). Walk the number line–An embod- ied training of numerical concepts. Trends in Neuroscience and Education, 2(2), 74–84.
Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. Child development, 1479-1498.
Linsey, J. S., Markman, A. B., & Wood, K. L. (2012). Design by analogy: A study of the WordTree method for problem re-representation. Journal of Mechanical Design, 134(4), 041009.
Lippa, R. A., Collaer, M. L., & Peters, M. (2010). Sex differences in mental rotation and line angle judgments are positively associated with gender equality and economic development across 53 nations. Archives of sexual behavior, 39, 990-997.
Lohman, D. F. (1996). Spatial ability and g. In I. Dennis & P. Tapsfield (Eds.), Human abilities: Their nature and measurement (pp. 97–116). Hillsdale, NJ: Erlbaum.
Looijenga, A., Klapwijk, R., & de Vries, M. J. (2015). The effect of iteration on the design performance of primary school children. International journal of technology and design education, 25, 1-23.
Lowrie, T., Logan, T., & Hegarty, M. (2019). The influence of spatial visualization training on students' spatial reasoning and mathematics performance. Journal of Cognition and Development, 20(5), 729-751.
Lowrie, T., Logan, T., & Ramful, A. (2016). Spatial Reasoning Influences Students' Performance on Mathematics Tasks. Mathematics Education Research Group of Australasia.
Lowrie, T., Logan, T., & Ramful, A. (2017). Visuospatial training improves elementary students' math- ematics performance. British Journal of Educational Psychology, 87(2), 170–186.
MacPhee, D., Farro, S., & Canetto, S. S. (2013). Academic self-efficacy and performance of underrepresented STEM majors: Gender, ethnic, and social class patterns. Analyses of Social Issues and Public Policy, 13(1), 347-369.
Maeda, Y., & Yoon, S. Y. (2013). A meta-analysis on gender differences in mental rotation ability measured by the Purdue spatial visualization tests: Visualization of rotations (PSVT: R). Educational Psychology Review, 25, 69-94.
Maresch, G. (2014). Strategies for assessing spatial ability tasks. Journal for Geometry and Graphics, 18(1), 125-132.
Maresch, G., & Sorby, S. A. (2021). Perspectives on spatial thinking. Journal for Geometry and Graphics, 25(2), 271-293.
Marshall, P. (2007). Do tangible interfaces enhance learning?. In Proceedings of the 1st international conference on Tangible and embedded interaction (pp. 163-170).
Mathewson, J. H. (1999). Visual-spatial thinking: An aspect of science overlooked by educators. Science Education, 83(1), 33–54.
Mayer, R. E. (Ed.). (2005). The Cambridge handbook of multimedia learning. Cambridge university press.
Mickelson, W. T., & Heaton, R. M. (2004). Primary teachers' statistical reasoning about

data. In The challenge of developing statistical literacy, reasoning and thinking (pp. 327-352). Dordrecht: Springer Netherlands.
Milner-Bolotin, M., & Nashon, S. M. (2012). The essence of student visual–spatial literacy and higher order thinking skills in undergraduate biology. Protoplasma, 249(1), 25–30.
Mix, K. S., Levine, S. C., Cheng, Y. L., Stockton, J. D., & Bower, C. (2021). Effects of spatial train- ing on mathematics in first and sixth grade children. Journal of Educational Psychology, 113(2), 304–314.
Mix, K. S., Levine, S. C., Cheng, Y. L., Young, C., Hambrick, D. Z., Ping, R., & Konstantopoulos, S. (2016). Separate but correlated: The latent structure of space and mathematics across development. Journal of Experimental Psychology: General, 145(9), 1206.
Moreno, D. P., Yang, M. C., Hernández, A. A., Linsey, J. S., & Wood, K. L. (2015). A step beyond to overcome design fixation: a design-by-analogy approach. Design computing and cognition'14 (pp. 607–624). Springer International Publishing.
Morrison, R. G., & Wallace, B. (2001). Imagery vividness, creativity and the visual arts. Journal of Mental Imagery, 25(3-4), 135–152.
Muller, W. (1989). Design discipline and the significance of visuo-spatial thinking. Design Studies, 10(1), 12–23.
Mustafa, N., Ismail, Z., Tasir, Z., & Mohamad Said, M. N. H. (2016). A meta-analysis on effective strategies for integrated STEM education. Advanced Science Letters, 22(12), 4225-4228.
Myford, C. M., & Wolfe, E. W. (2003). Detecting and measuring rater effects using many-facet Rasch measurement: Part I. Journal of applied measurement, 4(4), 386-422.
National Academy of Engineering & National Research Council. (2014). STEM integration in K-12 education: Status, prospects, and an agenda for research. Washington, DC: The National Academies Press.
National Research Council. (2006). Learning to think spatially: GIS as a support system in the K–12 curriculum. The National Academies Press.
National Research Council. (2011). Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics. Washington, DC: The National Academies Press.
Netwong, T. (2018). Development of problem solving skills by integration learning following stem education for higher education. International Journal of Information and Education Technology, 8(9), 639-643.
Neuburger, S., Ruthsatz, V., Jansen, P., & Quaiser-Pohl, C. (2015). Can girls think spatially? Influence of implicit gender stereotype activation and rotational axis on fourth graders' mental-rotation performance. Learning and Individual Differences, 37, 169-175.
Newcombe, N. S. (2010). Picture this: Increasing math and science learning by improving spatial thinking. American educator, 34(2), 29.
Newcombe, N. S. (2013). Seeing relationships: Using spatial thinking to teach science, mathematics, and social studies. American Educator, 37(1), 26–31 and 40.
Newcombe, N. (2017). Harnessing Spatial Thinking to Support STEM Learning. OECD Education Working Papers, No. 161. OECD Publishing.
Newcombe, N. S., & Frick, A. (2010). Early education for spatial intelligence: Why, what, and how. Mind, Brain, and Education, 4(3), 102–111.
Newcombe, N., & Huttenlocher, J. (2000). Making space: The development of spatial representation and reasoning. MIT Press.

Newcombe, N. S., & Shipley, T. F. (2015). Thinking About Spatial Thinking: New Typology, New Assessments. In: Gero, J. (eds) Studying Visual and Spatial Reasoning for Design Creativity. Springer, Dordrecht.
Newcombe, N. S., & Stieff, M. (2012). Six myths about spatial thinking. International Journal of Science Education, 34(6), 955–971.
Ng, O. L., & Chan, T. (2019). Learning as Making: Using 3D computer-aided design to enhance the learning of shape and space in STEM-integrated ways. British Journal of Educational Technology, 50(1), 294-308.
NGSS Lead States. (2013). Next generation science standards: For states, by states. The National Academies Press.
No More Marking. (n.d.). Accessed on May 1, 2024, from https://www.nomoremarking.com/
Noë, A. (2004). Action in perception. Cambridge, MA: MIT Press.
Novick, L. R., Hurley, S. M., & Francis, M. (1999). Evidence for abstract, schematic knowledge of three spatial diagram representations. Memory & Cognition, 27, 288–308.
Nusbaum, E. C., & Silvia, P. J. (2011). Are intelligence and creativity really so different? Fluid intelligence, executive processes, and strategy use in divergent thinking. Intelligence, 39, 36–45.
Nusbaum, E. C., Silvia, P. J., & Beaty, R. E. (2014). Ready, set, create: What instructing people to “be creative” reveals about the meaning and mechanisms of divergent thinking. Psychology of Aesthetics, Creativity, and the Arts, 8(4), 423–432.
Oberle, A. (2020). Advancing students' abilities through the geo-inquiry process. Journal of Geography, 119(2), 43-54.
Okuda, S. M., Runco, M. A., & Berger, D. E. (1991). Creativity and the finding and solving of real- world problems. Journal of Psychoeducational Assessment, 9(1), 45–53.
O'Malley, C., & Fraser, D. S. (2004). Literature review in learning with tangible technologies (Report 12). Bristol: Nesta FutureLab Series.
O'Quin, K., & Besemer, S. P. (1989). The development, reliability, and validity of the revised creative product semantic scale. Creativity Research Journal, 2(4), 267-278.
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.
Ormand, C. J., Shipley, T. F., Tikoff, B., Dutrow, B., Goodwin, L. B., Hickson, T., Atit, K., Gagnier, K., & Resnick, I. (2017). The Spatial Thinking Workbook: A research-validated spatial skills curriculum for geology majors. Journal of Geoscience Education, 65(4), 423–434.
Oxman, R. (2002). The thinking eye: visual re-cognition in design emergence. Design studies, 23(2), 135-164.
Pallasmaa, J. (2017). Embodied and existential wisdom in architecture: The thinking hand. Body & Society, 23(1), 96–111.
Palmiero, M., Nori, R., Aloisi, V., Ferrara, M., & Piccardi, L. (2015). Domain-specificity of creativity: A study on the relationship between visual creativity and visual mental imagery. Frontiers in Psychology, 6, 1870.
Papert, S. (1991). Situating constructionism. Cambridge, MA: MIT Press.
Papert, S. (1994). The children’s machine: rethinking school in the age of the computer Revised ed New York: Basic Books.
Paulus, P. B., Kohn, N. W., & Arditti, L. E. (2011). Effects of quantity and quality

instructions on brainstorming. The Journal of Creative Behavior, 45(1), 38-46.
Pedgley, O., & Şener, B. (2025). THE ROLE OF PHYSICAL SKETCH MODELLING IN THE DIGITAL DESIGN ERA. In DS 137: Proceedings of the International Conference on Engineering and Product Design Education (E&PDE 2025) (pp. 613-618).
Peng, A., & Sollervall, H. (2014). Primary School Students' Spatial Orientation Strategies in an Outdoor Learning Activity Supported by Mobile Technologies. Online Submission, 2(4), 246-256.
Perin, C. (2021). What students learn with personal data physicalization. IEEE Computer Graphics and Applications, 41(6), 48-58.
Piaget, J. (1952). The child's conception of number. New York: Humanities Press.
Pinheiro, J. C., & Bates, D. M. (2000). Linear mixed-effects models: basic concepts and examples. Mixed-effects models in S and S-Plus, 3-56.
Plucker, J. A. (1999). Is the proof in the pudding? Reanalyses of Torrance's (1958 to present) longitudinal data. Creativity research journal, 12(2), 103-114.
Pollitt, A. (2012). The method of adaptive comparative judgement. Assessment in Education: principles, policy & practice, 19(3), 281-300.
Prain, V., & Waldrip, B. (2010). Representing science literacies: An introduction. Research in science education, 40(1), 1-3.
Purcell, A. T., & Gero, J. S. (1996). Design and other types of fixation. Design Studies, 17(4), 363–383.
Purcell, A. T., & Gero, J. S. (1998). Drawings and the design process: A review of protocol studies in design and other disciplines and related research in cognitive psychology. Design Studies, 19(4), 389–430.
Qian, L., & Gero, J. S. (1996). Function–behavior–structure paths and their role in analogy-based design. AI EDAM, 10(4), 289–312.
Raat, J. H., De Vries, M. J. (Eds.). (1986). What do girls and boys think of technology?: Report PATT- workshop, March 6–11, 1986, Eindhoven University of Technology. Eindhoven University of Technology.
Raju, G., Sorby, S., & Reid, C. (2023). Exploring the Role of Spatial Visualization in Design Process of Undergraduate Engineering Students. Review directory American Society for Engineering Education.
Ramadas, J. (2009). Visual and spatial modes in science learning. International Journal of Science Education, 31(3), 301-318.
Ramey, K. E., Stevens, R., & Uttal, D. H. (2020). In-FUSE-ing STEAM learning with spatial reasoning: Distributed spatial sensemaking in school-based making activities. Journal of educational psychology, 112(3), 466.
Ramey, K. E., & Uttal, D. H. (2017). Making sense of space: Distributed spatial sensemaking in a middle school summer engineering camp. Journal of the Learning Sciences, 26(2), 277-319.
Ramful, A., Lowrie, T., & Logan, T. (2017). Measurement of spatial ability: Construction and validation of the spatial reasoning instrument for middle school students. Journal of Psychoeducational Assessment, 35(7), 709-727.
Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. Journal of experimental child psychology, 91(2), 137-157.
Reid, C., & Sorby, S. (2023). Exploring The Development Of Engineering Design Creativity And The Role Of Spatial Skills In This Process. European Society for Engineering Education (SEFI).
Ritz, J. M., & Fan, S. C. (2015). STEM and technology education: International state-of-

the-art. International Journal of Technology and Design Education, 25(4), 429-451.
Robitzsch, A. (2021). sirt: Supplementary Item Response Theory Models (Version R package version 3.10-118) [R]. Accessed on October 4, 2024, from https://CRAN.R-project.org/package=sirt
Rosenman, M. A., & Gero, J. S. (2013). Creativity in design using a design prototype approach. In Modeling creativity and knowledge-based creative design (pp. 111-138). Psychology Press.
Runco, M. A. (2010). Divergent thinking, creativity, and ideation. The Cambridge handbook of creativity, 413, 446.
Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. Creativity Research Journal, 24, 66–75.
Runco, M. A., & Charles, R. E. (1993). Judgments of originality and appropriateness as predictors of creativity. Personality and individual differences, 15(5), 537-546.
Runco, M. A., Nemiro, J., & Walberg, H. J. (1998). Personal explicit theories of creativity. The Journal of Creative Behavior, 32(1), 1-17.
Rustaman, N. Y., & Rahmat, A. (2017). Types of reasoning in framing based plant anatomy and it relation to spatial thinking. In Journal of Physics: Conference Series (Vol. 812, No. 1, p. 012055). IOP Publishing.
Rutland, M., & Barlex, D. (2008). Perspectives on pupil creativity in design and technology in the lower secondary curriculum in England. International Journal of Technology and Design Education, 18, 139-165.
Sadler, D. R. (2008). Transforming holistic assessment and grading into a vehicle for complex learning. In Assessment, learning and judgement in higher education (pp. 1-19). Dordrecht: Springer Netherlands.
Sanders, M. (2009). STEM, STEM education, STEMmania. The Technology Teacher, 68(4), 20–26.
Saretzki, J., Forthmann, B., Benedek, M., Andrae, R., Meinzer, N., Bartlmä, L., & Braun, F. (2024). A systematic quantitative review of divergent thinking assessments. Psychology of Aesthetics, Creativity and the Arts.
Sarkar, P., & Chakrabarti, A. (2011). Assessing design creativity. Design studies, 32(4), 348-383.
Saw, G., Chang, C. N., & Chan, H. Y. (2018). Cross-sectional and longitudinal disparities in STEM career aspirations at the intersection of gender, race/ethnicity, and socioeconomic status. Educational Researcher, 47(8), 525-531.
Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll Model of Intelligence. In D. P. Flanagan, & P. L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues (3rd ed., pp. 99–144). New York, NY: Guilford Press.
Schön, D. A. (1983). The reflective practicioner: how professionals think in action. New York: Basic Books.
Schön, D. A., & Wiggins, G. (1992). Kinds of seeing and their functions in designing. Design Studies, 13(2), 135–156.
Schut, A., Klapwijk, R., Gielen, M., van Doorn, F., & de Vries, M. (2020). Uncovering early indicators of fixation during the concept development stage of children’s design processes. International Journal of Technology and Design Education, 30, 951–972.
Search, P. (2015). Interactive multisensory data representation. In International Conference of Design, User Experience, and Usability (pp. 363-373). Cham: Springer International Publishing.
Seery, N., Buckley, J., Delahunty, T., & Canty, D. (2019). Integrating learners into the

assessment process using adaptive comparative judgement with an ipsative approach to identifying competence based gains relative to student ability levels. International Journal of Technology and Design Education, 29(4), 701-715.
Seitamaa-Hakkarainen, P., Kangas, K., Raunio, A. M., & Hakkarainen, K. (2012). Collaborative Design Practices in Technology Mediated Learning. Design and Technology Education, 17(1), 54-65.
Shaer, O., & Hornecker, E. (2010). Tangible user interfaces: past, present, and future directions. Foundations and Trends® in Human–Computer Interaction, 3(1–2), 4-137.
Shah, J. J., Smith, S. M., & Vargas-Hernandez, N. (2003). Metrics for measuring ideation effectiveness. Design studies, 24(2), 111-134.
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.
Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science, 171(3972), 701–703.
Shepard, R. N., & Feng, C. (1972). A chronometric study of mental paper folding. Cognitive Psychology, 3(2), 228–243.
Silvia, P. J., Beaty, R. E., & Nusbaum, E. C. (2013). Verbal fluency and creativity: General and specific contributions of broad retrieval ability (Gr) factors to divergent thinking. Intelligence, 41(5), 328–340.
Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: exploring the reliability and validity of new subjective scoring methods. Psychology of Aesthetics, Creativity, and the Arts, 2(2), 68.
Simeon, M. I., Samsudin, M. A., & Yakob, N. (2022). Effect of design thinking approach on students' achievement in some selected physics concepts in the context of STEM learning. International Journal of Technology and Design Education, 32(1), 185-212.
Simms, N. K., & Richland, L. E. (2019). Generating relations elicits a relational mindset in children. Cognitive Science, 43(10), e12795.
Sisman, B., Kucuk, S., & Yaman, Y. (2021). The effects of robotics training on children's spatial ability and attitude toward STEM. International Journal of Social Robotics, 13(2), 379-389.
Sluis-Thiescheffer, W., Bekker, T., Eggen, B., Vermeeren, A., & De Ridder, H. (2016). Measuring and comparing novelty for design solutions generated by young children through different design methods. Design Studies, 43, 48-73.
Smith, G. J., & Carlsson, I. (1983). Creativity in early and middle school years. International Journal of Behavioral Development, 6(2), 167-195.
Sonneveld, L., Klapwijk, R. M., & Stappers, P. J. (2024). Constructing and storytelling: accommodating different play orientations in learning spatial thinking. In Frontiers in Education (Vol. 9, p. 1307951). Frontiers Media SA.
Sorby, S. A. (2009). Educational research in developing 3-D spatial skills for engineering students. International Journal of Science Education, 31(3), 459-480.
Sorby, S. A., & Baartmans, B. J. (2000). The development and assessment of a course for enhancing the 3-D spatial visualization skills of first year engineering students. Journal of Engineering Education, 89(3), 301–307.
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.

Stables, K., & Kimbell, R. (2000). The unpickled portfolio: Pioneering performance assessment in design and technology. In Design and Technology International Millennium Conference, edited by Richard Kimbell, 195-202. London: The Design and Technology Association.
Stammes, H., Henze, I., Barendsen, E., & de Vries, M. (2020). Bringing design practices to chemistry classrooms: studying teachers' pedagogical ideas in the context of a professional learning community. International Journal of Science Education, 42(4), 526-546.
Starkey, E., Toh, C. A., & Miller, S. R. (2016). Abandoning creativity: The evolution of creative ideas in engineering design course projects. Design Studies, 47, 47-72.
Steedle, J. T., & Ferrara, S. (2016). Evaluating comparative judgment as an approach to essay scoring. Applied Measurement in Education, 29(3), 211-223.
Sternberg, R. J. (1977). Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Lawrence Erlbaum.
Sternberg, R. J., & Lubart, T. I. (1999). The concept of creativity: Prospects and paradigms. Handbook of creativity, 1(3-15).
Stevens, L., Kopnina, H., Mulder, K., & De Vries, M. (2021). Biomimicry design thinking education: A base-line exercise in preconceptions of biological analogies. International Journal of Technology and Design Education, 31, 797–814.
Stevens, L., Bidwell, D., Fehler, M., & Singhal, A. (2022). The art and science of biomimicry—abstracting design principles from nature. Transdisciplinarity (pp. 649–687). Springer International Publishing.
Stieff, M. (2011). Improving representational competence using molecular simulations embedded in inquiry activities. Journal of research in science teaching, 48(10), 1137-1158.
Stieff, M., Bateman, R. C., & Uttal, D. H. (2005). Teaching and learning with three-dimensional representations. Visualization in Science Education.
Stieff, M., Dixon, B. L., Ryu, M., Kumi, B. C., & Hegarty, M. (2014). Strategy training eliminates sex differences in spatial problem solving in a stem domain. Journal of Educational Psychology, 106(2), 390.
Stieff, M., Hegarty, M., & Dixon, B. (2010). Alternative strategies for spatial reasoning with diagrams. In International conference on theory and application of diagrams (pp. 115-127). Berlin, Heidelberg: Springer Berlin Heidelberg.
Stieff, M., Ryu, M., Dixon, B., & Hegarty, M. (2012). The role of spatial ability and strategy preference for spatial problem solving in organic chemistry. Journal of Chemical Education, 89(7), 854–859.
Stieff, M., & Uttal, D. (2015). How much can spatial training improve STEM achievement?. Educational Psychology Review, 27(4), 607-615.
Stigler, J. W. (1984). “Mental abacus”: The effect of abacus training on Chinese children's mental calculation. Cognitive Psychology, 16(2), 145-176.
Stoet, G., & Geary, D. C. (2018). The gender-equality paradox in science, technology, engineering, and mathematics education. Psychological science, 29(4), 581-593.
Stohlmann, M., Moore, T. J., & Roehrig, G. H. (2012). Considerations for teaching integrated STEM education. Journal of Pre-College Engineering Education Research (J-PEER), 2(1), 4.
Strimel, G. J., Bartholomew, S. R., Purzer, S., Zhang, L., & Ruesch, E. Y. (2021). Informing engineering design through adaptive comparative judgment. European Journal of Engineering Education, 46(2), 227-246.
Stull, A. T., Gainer, M. J., & Hegarty, M. (2018). Learning by enacting: The role of

embodiment in chemistry education. Learning and Instruction, 55, 80-92.
Suh, J., & Cho, J. Y. (2020). Linking spatial ability, spatial strategies, and spatial creativity: A step to clarify the fuzzy relationship between spatial ability and creativity. Thinking Skills and Creativity, 35, 100628.
Sutton, K., & Williams, A. (2010). Implications of Spatial Abilities on Design Thinking, in Durling, D., Bousbaci, R., Chen, L, Gauthier, P., Poldma, T., Roworth-Stokes, S. and Stolterman, E (eds.), Design and Complexity - DRS International Conference 2010, 7-9 July, Montreal, Canada.
Suwa, M., Gero, J., & Purcell, T. (2000). Unexpected discoveries and S-invention of design require- ments: Important vehicles for a design process. Design Studies, 21(6), 539–567.
Suwa, M., Purcell, T., & Gero, J. (1998). Macroscopic analysis of design processes based on a scheme for coding designers' cognitive actions. Design studies, 19(4), 455-483.
Suwa, M., & Tversky, B. (1997). What do architects and students perceive in their design sketches? A protocol analysis. Design studies, 18(4), 385–403.
Szechter, L. E., & Liben, L. S. (2004). Parental guidance in preschoolers' understanding of spatial-graphic representations. Child Development, 75(3), 869-885.
Tanaś, Ł., & Szarek, K. M. (2021). Beyond inhibitory control: Relationship between spatial and social skills in preschool children. Cognitive Development, 59, 101084.
Tassoul, M. (2009). Creative facilitation (3rd ed.). Delft: VSSD.
Taylor, H. A., & Hutton, A. (2013). Think3d!: Training spatial thinking fundamental to STEM education. Cognition and Instruction, 31(4), 434–455.
Taylor, A. S., Lindley, S., Regan, T., Sweeney, D., Vlachokyriakos, V., Grainger, L., & Lingel, J. (2015). Data-in-place: Thinking through the relations between data and community. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 2863-2872).
Tedjosaputro, M. A., Shih, Y.-T., Niblock, C., & Pradel, P. (2018). Interplay of sketches and mental imagery in the design ideation stage of novice designers. Design Journal, 21(1), 59–83.
The World Bank. (2018). What a waste global database. Data catalog. Accessed on January 23, 2023, from https://datacatalog.worldbank.org/search/dataset/0039597
Thurstone, L. L. (1955). Differential growth of mental abilities. In Science (Vol. 121, No. 3148, pp. 627). American Association for the Advancement of Science.
Tian, M., Luo, T., Ding, J., Wang, X., & Cheung, H. (2021). Spatial ability and theory of mind: A mediating role of visual perspective taking. Child Development, 92(4), 1590-1604.
Tillman, D. A., An, S. A., Cohen, J. D., Kjellstrom, W., & Boren, R. L. (2014). Exploring wind power: Improving mathematical thinking through digital fabrication. Journal of Educational Multimedia and Hypermedia, 23(4), 401-421.
Tsang, J. M., Blair, K. P., Bofferding, L., & Schwartz, D. L. (2015). Learning to "see" less than nothing: putting perceptual skills to work for learning numerical structure. Cognition and Instruction, 33(2), 154-197.
Tsomokos, D. I., & Flouri, E. (2025). From spatial to social competence: The association between spatial ability and prosocial behaviour in childhood. British Journal of Developmental Psychology.
Tomes, A., Oates, C., & Armstrong, P. (1998). Talking design: negotiating the verbal–visual translation. Design Studies, 19(2), 127-142.
Torrance, E. P. (1974). Torrance tests of creative thinking. Lexington, MA: Personnel Press.

Toth, A. J., & Campbell, M. J. (2019). Investigating sex differences, cognitive effort, strategy, and performance on a computerised version of the mental rotations test via eye tracking. Scientific Reports, 9(1), 19430.

Tracey, T. J. (2010). Relation of interest and self-efficacy occupational congruence and career choice certainty. Journal of vocational behavior, 76(3), 441-447.

Träff, U., Olsson, L., Skagerlund, K., Skagenholt, M., & Östergren, R. (2019). Logical reasoning, spatial processing, and verbal working memory: Longitudinal predictors of physics achievement at age 12–13 years. Frontiers in psychology, 10, 1929.

Tversky, B. (2001). "Spatial schemas in depictions" in Spatial schemas and abstract thought. ed. M. Gattis (Cambridge, MA: The MIT Press), 79–112.

Tversky, B. (2015). On Abstraction and Ambiguity. In: Gero, J. (eds) Studying Visual and Spatial Reasoning for Design Creativity. Springer, Dordrecht.

Tytler, R., & Hubber, P. (2016). Constructing Representations to Learn Science. In: Hand, B., McDermott, M., Prain, V. (eds) Using Multimodal Representations to Support Learning in the Science Classroom.

Tytler, R., & Prain, V. (2010). A framework for re-thinking learning in science from recent cognitive science perspectives. International Journal of Science Education, 32(15), 2055-2078.

Tytler, R., Prain, V., Aranda, G., Ferguson, J., & Gorur, R. (2020). Drawing to reason and learn in science. Journal of Research in Science Teaching, 57(2), 209-231.

Tytler, R., Prain, V., Hubber, P., & Haslam, F. (2013). Reasoning in science through representation. In Constructing representations to learn in science (pp. 83-107). Rotterdam: SensePublishers.

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.

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.

Uttal, D. H., McKee, K., Simms, N., Hegarty, M., & Newcombe, N. S. (2024). How can we best assess spatial skills? Practical and conceptual challenges. Journal of Intelligence, 12(1), 8.

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.

Vande Moere, A., & Patel, S. (2010). "The physical visualization of information: designing data sculptures in an educational context" in Visual information communication. eds. M. Huang, Q. Nguyen and K. Zhang (Boston, MA: Springer US), 1–23.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. Perceptual and motor skills, 47(2), 599-604.

Van Mechelen, M. (2016). Designing technologies for and with children: Theoretical reflections and a practical inquiry towards a co-design toolkit. Doctoral thesis.

Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. Contemporary Educational Psychology, 31(2), 142-166.

Vendetti, M., Matlen, B., Richland, L., & Bunge, S. (2015). Analogical reasoning in the classroom: Insights from cognitive science. Mind, Brain, and Education, 9(2), 100–106.

Verhavert, S., Bouwer, R., Donche, V., & De Maeyer, S. (2019). A meta-analysis on the reliability of comparative judgement. Assessment in Education: Principles, policy & practice, 26(5), 541-562.

Verhavert, S., Furlong, A., & Bouwer, R. (2022). The accuracy and efficiency of a reference-based adaptive selection algorithm for comparative judgment. Frontiers in Education, 6.

Visser, F. S., Stappers, P. J., Van der Lugt, R., & Sanders, E. B. (2005). Contextmapping: experiences from practice. CoDesign, 1(2), 119-149.

Voogt, J., & Roblin, N. P. (2012). A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies. Journal of curriculum studies, 44(3), 299-321.

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.

Waldschütz, H., & Hornecker, E. (2020). The importance of data curation for data physicalization. In Companion publication of the 2020 ACM designing interactive systems conference (pp. 293-297).

Wang, L., Cao, C., Zhou, X., & Qi, C. (2022). Spatial abilities associated with open math problem solving. Applied Cognitive Psychology, 36(2), 306-317.

Wang, M. T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. Educational psychology review, 29, 119-140.

Ware, C. (2004). Information visualization: Perception for design. Burlington, MA: Morgan Kaufmann.

Webb, R. M., Lubinski, D., & Benbow, C. P. (2007). Spatial ability: A neglected dimension in talent searches for intellectually precocious youth. Journal of Educational Psychology, 99(2), 397.

Weinberger, A. B., Iyer, H., & Green, A. E. (2016). Conscious augmentation of creative state enhances “real” creativity in open-ended analogical reasoning. PLoS ONE, 11(3), e0150773.

Wendell, K. B., & Rogers, C. (2013). Engineering design-based science, science content performance, and science attitudes in elementary school. Journal of Engineering Education, 102(4), 513-540.

Whitehouse, C. (2012). Testing the validity of judgements about geography essays using the Adaptive Comparative Judgement method. Manchester: AQA Centre for Education Research and Policy.

Willett, W., & Huron, S. (2016, October). A constructive classroom exercise for teaching infovis. In Pedagogy of Data Visualization Workshop at IEEE VIS 2016.

Williams, C. B., Gero, J., Lee, Y., & Paretti, M. (2010). Exploring spatial reasoning ability and design cognition in undergraduate engineering students. In International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (Vol. 44144, pp. 669– 676).

Wilson, M. (2002). Six views of embodied cognition. Psychonomic bulletin & review, 9(4), 625-636.

Wilson, V., & Harris, M. (2004). Creating change? A review of the impact of design and technology in schools in England. Journal of Technology Education, 15(2), 6–18.

Wilson, J. O., Rosen, D., Nelson, B. A., & Yen, J. (2010). The effects of biological examples in idea generation. Design Studies, 31(2), 169–186.

Wu, H. K., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning.

Science education, 88(3), 465-492.
Wun, T., Payne, J., Huron, S., & Carpendale, S. (2016). Comparing bar chart authoring with microsoft excel and tangible tiles. In Computer Graphics Forum (Vol. 35, No. 3, pp. 111-120).
Xu, C., & LeFevre, J.-A. (2016). Training young children on sequential relations among numbers and spatial decomposition: Differential transfer to number line and mental transformation tasks. Developmental Psychology, 52(6), 854–866.
Xu, M., Williams, P. J., Gu, J., & Zhang, H. (2020). Hotspots and trends of technology education in the International Journal of Technology and Design Education: 2000–2018. International Journal of Technology and Design Education, 30(2), 207-224.
Yuan, L., Uttal, D., & Gentner, D. (2017). Analogical processes in children's understanding of spatial representations. Developmental Psychology, 53(6), 1098.
Zahner, D., Nickerson, J. V., Tversky, B., Corter, J. E., & Ma, J. (2010). A fix for fixation? Rerepresenting and abstracting as creative processes in the design of information systems. AI EDAM, 24(2), 231–244.
Zhan, Z., Yao, X., & Li, T. (2022). Effects of association interventions on students' creative thinking, aptitude, empathy, and design scheme in a STEAM course: Considering remote and close association. International Journal of Technology and Design Education.
Zhao, J., & Vande Moere, A. (2008). Embodiment in data sculpture: a model of the physical visualization of information. In Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts (pp. 343–350).
Zhu, C. & Klapwijk, R. (2024). Assessing pupils' creative design ideas and prototypes using comparative judgment. 41st Pupils' Attitudes Towards Technology (PATT) Education Research Conference in Nanjing, China (pp. 290-296).
Zhu, C., Klapwijk, R., Silva-Ordaz, M., Spandaw, J., & de Vries, M. J. (2023). Cognitive and embodied mapping of data: an examination of children's spatial thinking in data physicalization. In Frontiers in Education (Vol. 8, p. 1308117). Frontiers Media SA.
Zhu, C., Klapwijk, R., Silva-Ordaz, M., Spandaw, J., & de Vries, M. J. (2024). Investigating the role of spatial thinking in children's design ideation through an open-ended design-by-analogy challenge. International Journal of Technology and Design Education, 34(5), 1733-1762.
Zhu, C., Leung, C. O. Y., Lagoudaki, E., Velho, M., Segura-Caballero, N., Jolles, D., Duffy, G., Maresch, G., Pagkratidou, M., & Klapwijk, R. (2023). Fostering spatial ability development in and for authentic STEM learning. Frontiers in Education, 8, 1138607.
Zhu, W., Shang, S., Jiang, W., Pei, M., & Su, Y. (2019). Convergent thinking moderates the relationship between divergent thinking and scientific creativity. Creativity Research Journal, 31(3), 320–328.

Appendices

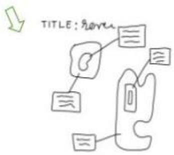
Appendix 1 Snacksafe on the Beach design task and booklet[7]

1 What features do you think your design needs to have to solve this design problem well?

2 Use this space to jot down more design ideas, try to think of unusual, original ideas

3 Take a look at all your ideas, circle your most unusual, original idea that could solve the design problem

4 Use this space to make detailed sketches of your selected idea.
What will it finally look like?
Sketch and explain on how your design can allow people to enjoy their fries on the beach without worrying about the seagulls. Make sure you add labels and notes to help others understand your design like this.



Title of your design:____

5 How will this design work?

[7] This task is designed to be printed on A3 paper and folded into an A5 booklet.

Appendix 2 Codebook—Design judges' comments on pupils' design ideation quality

Coding guidelines
- Each sentence is treated as the smallest unit of analysis and may receive multiple codes. One piece of comment may receive the same code more than once if that code appears in multiple sentences within the comment
- The (+/-) notation indicates that the category includes instances where the criterion is expressed in either positive or negative terms
- Codes should be assigned based on what judges explicitly mentioned in their comments, not on the researchers' interpretation of pupils' designs beyond what is directly mentioned in the comments

Code system	Definition	Examples
Novelty		
Original/unique (+/-)	Judges commenting on a design being innovative, new, original, unique, rare, fresh, not often seen, or proposing a new concept; also on a design being an obvious idea, similar to other frequently seen or existing solutions	"Ideas on the left are a bit boring because it is obvious" "one on the right mainly has one idea but that one is very unique and creative to me"
Different starting point (+)	Judges commenting on a design taking a different direction or perspective, changing the context, or reframing the scope of the problem; could be expressed in these forms: "instead of (a common way) the child did (new and different) way" "it's not about a (common idea)...but about…"	"the child did not think about the fries or the packaging but rather what the seagulls really want which in the end is just food" "The one on the right instead of being a container for fries is a container for fans to keep the gulls away which is creative in looking at the problem"
Creatively combining different ideas (+)	Judges commenting on a design being a creative and meaningful combination of different concepts that may have otherwise been common ideas	"Left has not so creative ideas but combines it to a somewhat creative end result" "the shining sticker is more creative and it is combined with a logo"
Uncommon mechanism (mechanical/structural) (+)	Judges commenting on an uncommon mechanism—a specific structure or mechanical component—in the design, including the shape, the structure, the compartments and components, the way of assembling/attaching parts of the design, or the motion or interaction generated due to these compartments, that are unusual, novel, or not seen before	"The idea of linking all these elements and make them activate in sequence is really unique" "the right one is a bit more creative with the mechanism of the box and the use of drones to deliver the fries"

Idea qualities		
Good (general) (+)	Judges commenting on an idea/design as good in general, incredible, awesome, amazing, nice, or cool	“Hard to pick one or the other because they followed a good design process” “right develops them further into newish novel ideas! Awesome!”
Smart (+)	Judges commenting on an idea/design as smart, clever, genius, brilliant	“the approach of connecting the predator with a piece of art (genius!)” “left side is a very elegant and smart solution to the problem”
Imagination (+)	Judges commenting on an idea/design that reflects imagination	“I like how the playfulness in the ideas and use of imagination” “making their brainstorm more creative in terms of senses imagined”
Interesting (+)	Judges commenting on an idea/design as interesting	“left has more interesting ideas like taming the seagulls” “left has ideas like scarecrow which is interesting way of using learning from one to another”
Fun & playful (+)	Judges commenting on an idea/design as fun, funny, or playful	“I like how the idea thinks of not only protecting the fries but also makes it fun by making a play park for children” “Funny they have written names on the fries”
Idea potential (+)	Judges commenting on how the ideas can have creative potentials for further development	“The right one have a lot of ideas about the fries box & mechanism which in the future they can scope down or develop them” “left has simple solution which makes it interesting and promising both”

Under-developed ideas (-)	Poor ideation quality, straightforward or surface-level shallow thinking, lack of effort, or when judges specifically mention there's only one idea with little elaboration, simple in a bad way, little explanation given, showing little-to-no development in ideas; this code is only negative	"Left is hard to understand because there is no annotation but it seems to be just one idea" "ideas on the right are really basic"
Aesthetics & desirability (+/-)	Judges commenting on the design being aesthetically pleasing, or that certain features of the design make it appealing and desirable for people	"the ideas have a nice look making them also appealing the eye" "The left one looks more fancy so I vote for that"
Sustainability (+/-)	Judges' appreciation or concern for whether the resources are reusable or not, or whether the solution appears to be sustainable, eco-friendly, recyclable or not	"the left considers problems like environmental protection" "it's also reduce food waste at the same time"
Usability		
Multiple elements/features (+/-)	Judges describing a design as containing multiple elements, compartments, multiple steps or procedures, multiple different features; (-) when there is a lack of multiple elements	"LEFT sticks to the packaging itself and explores ideas around that and involves different elements" "the right ideas are more creative as the child did not only think about packaging but did consider the use of sound smell as well as color"
Multiple functions/purposes (+/-)	Judges describing a design as performing multiple functions or serving different purposes; (-) when there is a lack of multiple functions	"The idea has multiple functions or solves more than one problem" "the different functions that the product integrates are all interconnected"
Useful & functional (+/-)	Judges commenting on whether a design/solution is functional, effective, useful, or whether the proposed solution can actually work to solve the problem or not; also commenting on the design being practical or convenient for the design problem or not	"mask and invisible fries mirror all the ideas "hide" and fries very well" "right doesn't provide many details about the product how it would work and leaves me wondering if it is really effective."

Simple & intuitive (+/-)	Judges commenting on a design product or solution as easy or intuitive to use; also refers to simplicity (elegance) of a solution as a desired trait; (-) when judges mention a design product or solution being unnecessarily complex or even chaotic	"sometimes using simplicity and good storytelling can make the difference even if the ides is not so mindblowing" "The design on the left looks a bit confusing and overcomplicated"
Tailored for the context (+)	Judges commenting on specific considerations demonstrated in the design product or solution that are tailored for the context, e.g., the design showed considerations for how it fit into the beach environment, or specifically the sandy/windy beach; or the design smartly leverage things that already exist on the beach, e.g., beach umbrella, beach bench	"the decision is based on how well they are integrated with the context of use" "I like the playfulness in the ideas and how fry smell blower and a fries shooter fits interestingly in the beach environment"
User experience > judges' considerations (+/-)	Judges expressing appreciation or concerns for the kinds of user experience indicated by pupils' design; could be possible adverse side effects that pupils have not thought of themselves (e.g., "people might not like being in the nets")	"the concept is a bit shallow because it doesn't explore how the device would affect the experience of eating fries" "most of the design ideas are easy to use and can be used by everyone"
User experience > pupils' considerations (+)	Judges noticing that the design reflected pupils' consideration of user experience, including making something easier to use for the users, adding specified features for specific groups, or pupils considering how both human and the seagulls could be the users of the design and how they would experience or interact with the design	"Ideas on the right keep different types of target groups in mind" "I love that the child did also think about people with disabilities and included that perspective in their ideas"
Feasibility		
Considering materials (+/-)	Judges commenting on pupils' consideration of the materials they need for making or producing their design	"but right additionally considers materiality and communicates how it wont hurt seaguls" "LEFT offers more variety and explores more the idea with different materials (raft net tent etc.)"

Involving technology (+/-)	Judges commenting on technology used or technical element involved in the design seen, could be positive or negative	"although they include some interesting technology such as sensors to detect seagulls they don't stand out" "Left: Use technology to feed the seagull (drone) also provide alternative idea like ultrasonic sound"
Cost-effectiveness (+/-)	Judges expressing appreciation or concern for the cost and resources induced by making or implement the proposed design, or judges noticing pupils' cost-effectiveness considerations	"the right one used touch-ID which is higher cost and unnecessary here" "the left one use extra bread to distract seagulls which costs more resources"
Realistic to make (+/-)	Judges commenting on whether it is possible/feasible to realize or realistic to make the design; (-) concern for seemingly unrealistic and impossible setups	"the one on the right is more creative but less plausible" "right is more appropriate for the given challenge since the size of the project is way smaller and easier to realize"
Presentation		
Elaborated details (+/-)	Judges mentioning that the presentation of the design is rich with elaborations and details or not	"each concept has more detailed explanations" "A bit difficult to assess the left one since not many explanations were given"
Clarity in explanation (+/-)	Judges commenting on the design presented being clearly or unclearly explained, in terms of clarity in both textual and figural explanations; also coded when judges find the ideas understandable or confusing	"None of the ideas is explained or illustrated clearly" "some drawings are a bit vague"
Quality of drawing (+/-)	Judges commenting on elaborative, aesthetically good, well-made drawing or poor drawing, or drawing showing the design from different viewpoints or angles (pure drawing quality, not to be mixed with clarity in drawing)	"the drawing is a bit too simple" "Left showed more visualization skills"
Storytelling (+)	Judges commenting on the design/sets of ideas showing good storytelling abilities, or presented with story plots	"right seems to have a small story which showcases the idea" "sometimes using simplicity and good storytelling can make the difference"

Problem-solving		
Thought-through solutions (+/-)	Judges explicit commenting on the design solution as thought- through, in-depth, worked-out, thorough, showing a range of considerations, a lot of thoughts, logical, coherent, cohesive, or not; it could involve noticing that the child addressed the multiple steps needed to implement the design, different possible scenarios that may be generated, or presenting the solutions from multiple different angles, or that the child took into consideration various possible issues arising from the design; only mentioning "elaborated" or "detailed" do not suffice this code and should go to "presentation > elaborated details"	"Right explores 2 more out of the box ideas in more depth." "the candidate thinks it trough towards a details newish packing design"
Meeting the design brief (+/-)	Judges consideration of whether the design solution is appropriate for the design brief (not harming the seagulls, appropriate for human users, appropriate for the beach environment); also when judges mention whether the design solution is problem-oriented, achieving the key design goal or not	"The design on the right is more appropriate for the given challenge" "Right has pulled it a bit too far with a chillroom with airconditioning, too far away from the assignment"
Idea generation		
Diverse directions (+/-)	Judges commenting on the brainstormed ideas being diverse, varied, or a broad, wide range of ideas that are of different kinds and in multiple directions, showing divergent thinking, and showing that the kid has explored the solution space widely; or judges mentioning that the ideas are focused, narrowed, all in similar direction, and did not explore the solution space much; the mentioning of "many different ideas" would be coded only as diverse directions and not quanities of ideas	"Ideas on the left are all in similar direction." "right offer a much wider and crazier exploration with no connection to each other"
Variations of a key idea (+/-)	Judges commenting on the brainstormed ideas being varied versions of a key idea or key concept, without going into different directions; also use this code when iteration on a key idea is seen in the brainstorm	"Ideas on the left are iterated nicely" "Right sticks to the same container but with different toppings"
Quantities of ideas (+/-)	Judges commenting on the quantities of ideas, such as, more, multiple, several, a lot of, multiple, a good amount of ideas, or, e.g., mentioning that there are fewer ideas or only two or three ideas; this code is given when the judge mentions only quantity of the ideas without mentioning whether the ideas are diverse or varied	"I only vote for left because it has an higher quantity of ideas" "right shows more ideas, quantity will lead to quality during a brainstorm"

Appendix 3 Codebook—Design judges' comments on pupils' design prototype quality

Coding guidelines
- Each sentence is treated as the smallest unit of analysis and may receive multiple codes. One piece of comment may receive the same code more than once if that code appears in multiple sentences within the comment
- The (+/-) notation indicates that the category includes instances where the criterion is expressed in either positive or negative terms
- Codes should be assigned based on what judges explicitly mentioned in their comments, not on the researchers' interpretation of pupils' designs beyond what is directly mentioned in the comments

Code system	Definition	Examples
Novelty		
Original/unique (+/-)	Judges commenting on a design being innovative, new, original, unique, rare, fresh, not often seen, or proposing a new concept; also on a design being an obvious idea, similar to other frequently seen or existing solutions	"It stands out as an original idea compared to the more straightforward container solution" "I did see this also in other ideas already though"
Different starting point (+)	Judges commenting on a design taking a different direction or perspective, changing the context, or reframing the scope of the problem; could be expressed in these forms: "instead of (a common way) the child did (new and different) way" "it's not about a (common idea)...but about…"	"The child focuses on the delivery of the fries and not the packaging" "it creates opportunities for exercising and meeting other people so changes the context a lot!"
Modify an otherwise usual idea (+)	Judges commenting on the core design idea being common, but noting that the creative features added to the design made it distinct from the otherwise conventional types of usage	"they also though of the bottle spraying based on a timer, which would make it a bit more innovative then already existing scent sprays" "I also like that the rethought how the fries can be stored in a bag rather than a box. Which is why I think it is the more creative option"
Creatively combining different ideas (+)	Judges commenting on a design being a creative and meaningful combination of different concepts that may have otherwise been common ideas	"I like that the left one uses play predators in combination with the sound of playing kids" "I think the left design is a bit more out-of- the-box thinking with the use of water and sound to repel the seagulls"

Uncommon mechanism (mechanical/structural) (+)	Judges commenting on an uncommon mechanism—a specific structure or mechanical component—in the design, including the shape, the structure, the compartments and components, the way of assembling/attaching parts of the design, or the motion or interaction generated due to these compartments, that are unusual, novel, or not seen before	"left design is more innovative using rubber mouth-like openings so it can keep original state automatically" "materials used are creative (mirroring string foldable technique)"
Creative/out-of-the-box (general) (+/-)	Judges commenting on a design being 'creative' or 'out-of-the-box' in a general way that does not fall into the above-mentioned categories, that is, without mentioning that it is new and rare, exhibiting uncommon mechanisms, combining different ideas, taking a different direction, or modifying a usual idea	"the drawing of the seagull is a creative idea" "in the end I think the left one is a bit more outside of the box thinking"
Idea qualities		
Good (general) (+)	Judges commenting on an idea/design as good in general, incredible, awesome, amazing, nice, or cool	"left one is a amazing idea for the packaging" "right seems like a cool idea"
Smart (+)	Judges commenting on an idea/design as smart, clever, genius, brilliant	"scaring the seagulls by using the mirror is quite smart"
Surprising (+)	Judges commenting on an idea/design as surprising, shocking, unexpected	"The left design surprised me and put a smile on my face"
Interesting (+)	Judges commenting on an idea/design as interesting	"the context aware solution is interesting and creative"
Fun & playful (+)	Judges commenting on an idea/design as fun, funny, or playful	"There is a mechanism that takes multiple steps to protect the fries in a funny way" "The kid thought about how to scare away seagulls in a playful way"
Considerate (+)	Judges commenting on an idea/design as thoughtful or considerate	"This one is more thoughtful and think about most detail" "The description for left one is very considerate and like a designer"
Idea potential (+)	Judges reasoning about how the idea/design makes them think further about certain aspects, designs having creative potential, or judges' reasoning that the design could have certain potential if improved	"the bubble idea might ruin the fun at the beach but it shows more opportunities to iterate" "another umbrella but creative starting point for further ideas"

Under-developed ideas (-)	Judges commenting on the design being poor, boring, a lack of in-depth thoughts, straightforward thinking, simple in a bad way, zero explanation given, showing little-to-no development in ideas; this code is only negative	"Though less is more left is too simple" "The right one does not really have a design its a tent in the form of a cube"
Aesthetics & desirability (+/-)	Judges commenting on the design being aesthetically pleasing, or that certain features of the design can make it appealing and desirable for people	"from an aesthetic perspective the right one looks better" "I like that they made it attractive for children with the seagull head"
Sustainability (+/-)	Judges' appreciation or concern for the whether the resources are reusable or not, or whether the solution appears to be sustainable, eco-friendly, recyclable or not	"New simple material (rubber plastic) seems reusable" "They also thought of things like energy conservation"
Explorative (+/-)	Judges commenting on whether the child has explored different possibilities, or considered expanding the solution space	"The right design is more explorative of different solution spaces" "But more ideas on how to safeguard the fries seem to have been explored"
Usability		
Multiple elements/features (+/-)	Judges describing a design as containing multiple elements, compartments, multiple steps or procedures, multiple different features; (-) when there is a lack of multiple elements	"There are multiple elements involved in the design" "The right has many features but they are all rather common"
Multiple functions/purposes (+/-)	Judges describing a design as performing multiple functions or serving different purposes; (-) when there is a lack of multiple functions	"the left one is the more creative design as the idea can be reused for many purposes" "It also uses learning from one kind of usage (scarecrow in farming) to another (eating fries) - which makes it interesting"
Useful & functional (+/-)	Judges commenting on whether a design/ solution is functional, effective, useful, or whether the proposed solution can actually work to solve the problem or not; also commenting on the design being practical or convenient for the design problem or not	"from the aspect of usage left is better the holes in right one might not easy letting the fries to go out" "the food gun design with the speaker is good but might be less convenient to use"

Simple & intuitive (+/-)	Judges commenting on a design product or solution as easy or intuitive to use; also refers to simplicity (elegance) of a solution as a desired trait; (-) also coded when judges mention a design product or solution being unnecessarily complex or even chaotic	"It is a pretty simple solution however still very effective" "Hard to understand and overcomplicated and focused on technological aspects"
Tailored for the context (+)	Judges mentioning that there are specific considerations demonstrated in the design product or solution that are tailored for the context, e.g., the design showed considerations for how it fit into the beach environment, or specifically the sandy/windy beach; or the design smartly leverage things that already exist on the beach, e.g., beach umbrella, beach bench	"the right is better fitting to the scene" "I like the way of thinking about the folding working with what is already there"
Customization (+)	Judges mentioning that the design offers customizable, personalized features or options, e.g., offering different sizes, different versions for different target groups or different scenarios	"right- considerate in providing different sizes" "left - that it is customizable and shows the both the front and side of the product making it more clear"
User experience > judges' considerations (+/-)	Judges expressing appreciation or concerns for the kinds of user experience indicated by pupils' design; could be possible adverse side effects that pupils have not thought of themselves (e.g., "people might not like being in the nets")	"the left one is using small cabinet house to protect but it might conflict with the idea that people want to enjoy the environment" "the fact that it's foldable makes it a even more enjoyable solution"
User experience > pupils' considerations (+)	Judges noticing that the design reflected pupils' consideration of user experience, including making something easier to use for the users, adding specified features for specific groups, or pupils considering how both human and the seagulls could be the users of the design and how they would experience or interact with the design	"Left idea is more creative since it thinks about the target group's comfort" "Left is more creative as it also covers something positive for the birds"
Feasibility		
Considering materials (+/-)	Judges commenting on pupils' consideration of the materials they need for making or producing their design	"Left also thinks about materials that should be used" "Some details are also given (size of the seagull predator the movement and materials chosen)"

Involving technology (+/-)	Judges commenting on technology used or technical element involved in the design seen, could be positive or negative	“It is nice that this person thought about the solar panel” “Details given are creative (".. specific volume chosen" "..run on batteries.." "..via bluetooth..")”
Cost-effectiveness (+/-)	Judges expressing appreciation or concern for the cost and resources induced by making or implement the proposed design, or judges noticing pupils’ cost-effectiveness considerations	"’Everyone will buy their own and buy refills of fries’ shows thinking of how the design is financially feasible” “The glass barrier seems a bit hard to afford and implement in every beach”
Realistic to make (+/-)	Judges commenting on whether it is possible/feasible to realize or realistic to make the design; (-) could involve concern for seemingly unrealistic and impossible setups	“The left design is more "Wizard-of-Oz" style there's quite some magic needed to make it work” “The one on the right instead is less creative but is a plausible solution”
Presentation		
Elaborated details (+/-)	Judges mentioning that the presentation of the design is rich with elaborations and details or not	“the portable fry box is well designed with options for 3 sauce dipping and detailed on how it work” “The solution on the right is way less elaborate and is based on a not very realistic technology”
Clarity in explanation (+/-)	Judges commenting on the design presented being clearly or unclearly explained, in terms of clarity in both textual and figural explanations; also coded when judges find the ideas understandable or confusing	“Simple yet also very complicated solution which isn't thought through explanation lacks.” “The design process in clear and notes are easy to understand”
Quality of drawing (+/-)	Judges commenting on elaborative, aesthetically good, well-made drawing or poor drawing, or drawing showing the design from different viewpoints or angles (pure drawing quality, not to be mixed with clarity in drawing)	“Nice drawings from multiple angle” “At least the drawing is quiet nice with the checkered design”

Problem-solving		
Thought-through solutions (+/-)	Judges explicit commenting on the design solution as thought- through, in-depth, worked-out, thorough, showing a range of considerations, a lot of thoughts, logical, coherent, cohesive, or not; it could involve noticing that the child addressed the multiple steps needed to implement the design, different possible scenarios that may be generated, or presenting the solutions from multiple different angles, or that the child took into consideration various possible issues arising from the design; mentions of "elaborated" or "detailed" alone are insufficient for this code and should be assigned to "presentation > elaborated details"	"Right idea is more creative because the kid worked out a full story / scenario" "the design does not seem thought in depth"
Meeting the design brief (+/-)	Judges consideration of whether the design solution is appropriate for the design brief (not harming the seagulls, appropriate for human users, appropriate for the beach environment); also when judges mention whether the design solution is problem-oriented, achieving the key design goal or not	"The right one seagull attractant uses chemical solutions to keep seagulls away from people but this might be against the design principle of not harming the seagull" "Creating a new safe space (= building in this sense) seems less creative than designing a very problem oriented box for the fries"

Acknowledgements

My PhD journey has been shaped by learning, exploration, collaboration, and trust. I am deeply thankful to the many people who supported me, challenged me, and stood by me throughout this process.

Thank you, Remke, my copromotor, for your insights on research and for the way you bring people together. You are passionate, creative, and always authentically yourself. I still remember overhearing you say, in my first year of the PhD, that I am both analytical and creative as a researcher. Surrounded by designers and makers in Delft, I did not see myself as particularly creative then. But your words nudged me to reflect on how I understand creativity and how I see myself; in a way, they later connected beautifully with my own research. I have come to realize that creativity is not only about thinking “outside the box,” but about approaching complex problems with curiosity and rigor, seeing them from new angles, and having the courage to explore open questions.

Thank you, Marc, my promotor, for sharing your extensive experience in research and publication, for your incisive feedback, and for introducing me to the PATT community—a vibrant and supportive community dedicated to technology and design education that you have led for more than 40 years. Thank you, Jeroen, my second supervisor, for bringing both the mathematician’s and mathematics educator’s perspectives into our progress meetings, and for your many thoughtful and critical questions that continually helped strengthen my research. Thank you, Jeff, for involving me in such interesting investigations and for your kind encouragement. I am inspired by how you excel in your many roles while remaining deeply committed to supporting your students. Thank you, Dietsje—I enjoyed the opportunity to work together with you on the final study of my PhD. Your insightful and thorough feedback shaped what began as a vague idea during the SellSTEM training at the Wolfgangsee into Chapter 5 of this thesis.

I am grateful to all my colleagues at SEEd for the diverse expertise you bring to our section. Thank you, Eveline and Leonie, for your valuable designerly insights into my research and for the many brainstorming sessions we shared, which ended not only with ideas but also with laughter. Thank you, Anna, for your thoughtful leadership and for the supportive environment you create. Thank you, Fiona, for all the help you provided and for always handling our requests efficiently. A special thanks to Remke, Marc, Hanno, Caroline, Peter, Eveline, Anna, Jeroen, Suzanne, Rohit, Megan, Sebastiaan, and Anne for your generous feedback and support during my mock defense.

Thank you to all the school pupils who participated in my research. I was constantly inspired by your creativity, your unique perspectives, and your curiosity about the world. I am also grateful to all the teachers I worked with from international schools across the Netherlands, especially Miros, Joel, Ana, Peter, Masa, Louise, Barbara, and Jesse, for generously opening your classrooms to me and for sharing the insights you gained through years of teaching and growing alongside your students. A special thank you to Miros—your passion for design and technology, along with the positivity you bring to your students, inspired me from the very start of my PhD journey. In addition, I would like to thank all the reviewers who kindly dedicated their time and thoughtful feedback to helping me shape my manuscripts.

Thanks to everyone in SellSTEM. I cherish the memories we shared during our training in Delft, Dublin, Salzburg, and Riga, and I truly appreciate the different perspectives, cultures, and knowledge that each of you brought to the team. Through the challenges we faced and the inside humor we shared, I came to know each of you not only as researchers but as individuals. A special thanks to Rohit for the countless lively discussions and the many to-do lists we never quite finished but kept each other accountable for. We learned from and supported each other throughout our PhDs, and I wish you every success as an inspiring educator, bringing your research into practice.

To my dear friends in China, the United States, the Netherlands, and elsewhere around the world, thank you for always being by my side—physically or virtually. I feel incredibly lucky to share this journey with you and to witness each of you finding your own path. In many ways, our lives feel as though they unfold in parallel, and through you, I see the great possibilities that life can hold. A special thanks to Ruicen, Shuyi, Susu, Xiyue, Reeves, Sally, Julian, Xinyi, Shengran, Yijing, Winnie, Jin, Wenyi, Yuan, Estefania, Preeti, Gayeon, Hongjin, and Ruiyuan for all the laughter and tears we shared over these years. Thank you to all the friends and coaches I met through CrossFit, bouldering, and at X. You inspire me to be resilient and persistent, to show up even on the hard days, and to remember that progress happens just beyond comfort.

Thank you, Gilles, my boyfriend, for your unwavering support, your kind heart, and your brilliant mind. You offer fresh perspectives on my research, experiment with delicious meals, and train together with me as we grow physically and mentally stronger. With your passion and dedication, I have no doubt that you will become a researcher who is committed to making a meaningful impact on society and inspiring future students. I am also deeply grateful to everyone in your family, especially Marleen, Stephan, Nils, and Jess, for every lovely holiday season we have spent together and for always making me feel at home in the Netherlands.

Finally, I would like to express my gratitude, in Chinese, to all my family members, teachers, and mentors who have supported and cared for me along the way. Your messages across time zones have always filled my heart with strength and happiness. I would like to thank my parents, Yinghua and Meihua, for their unconditional support and trust throughout my studies and PhD journey. You have always believed in every choice I made and every path I chose to follow. The positivity and perseverance I learned from both of you supported me through countless challenges and enabled me to give my best to everything I choose to do. I will continue forward, navigating life with your love and confidence firmly by my side.

最后，我想用中文向每一位关心支持我的亲人及师长表达由衷的感谢。你们跨越时区的问候，总能让我的心中充满力量与幸福。感谢我的爸爸妈妈，你们一直以来给予我无条件的支持与信任，是我最坚实的后盾和最温暖的港湾。这本博士论文集凝结着自小以来你们言传身教所给予我的积极与坚韧：选择做一件事，便全力以赴将其做好。未来的道路上，我会带着你们的爱与这份笃定，从容而坚定地走好每一步。

Curriculum Vitae

Caiwei Zhu 朱彩维

2025-	Postdoctoral Researcher Contributing to an NRO-funded project on the use of virtual reality (VR) to enhance learning in medical education, the reuse and integration of VR applications across educational contexts, and the underlying mechanisms of immersive learning informed by the Cognitive Affective Model of Immersive Learning	Instructional Technology section, Department of Learning, Data analytics, and Technology, University of Twente
2021 - 2025	PhD Candidate Early-stage researcher in the Marie Skłodowska-Curie Innovative Training Network project SellSTEM (Spatially Enhanced Learning Linked to STEM), examining primary and secondary school pupils' use of spatial thinking through design-based learning projects and investigating the relationship between design, creativity, and spatial ability	Science & Engineering Education section (SEEd), Faculty of Applied Sciences, Delft University of Technology
2020 - 2021	Master of Education in Human Development & Psychology	Harvard University, USA
2018 - 2020	Bachelor of Science in Human Development (with distinction)	Cornell University, USA
2017 - 2018	Bachelor Studies	University of Virginia, USA
2011 - 2017	Secondary School	Shenzhen Foreign Languages School, China

List of Publications

Zhu, C., Buckley, J., Klapwijk, R., Spandaw, J., de Vries, M. J. (2025). A holistic look at creativity: Evaluating pupils' creative design ideation and prototypes through comparative judgment. International Journal of Technology and Design Education.

Zhu, C., Klapwijk, R., Silva-Ordaz, M., Spandaw, J., de Vries, M. J. (2024). Investigating the role of spatial thinking in children's design ideation through an open-ended design-by-analogy challenge. International Journal of Technology and Design Education.

Zhu, C., Klapwijk, R., Silva-Ordaz, M., Spandaw, J., de Vries, M. J. (2023). Cognitive and embodied mapping of data: an examination of children's spatial thinking in data physicalization. Frontiers in Education.

Zhu, C., Leung, C. O-Y., Lagoudaki, E., Velho, M., Segura-Caballero, N., Jolles, D., Duffy, G., Maresch, G., Pagkratidou, M., & Klapwijk, R. (2023). Fostering spatial ability development in and for authentic STEM learning. Frontiers in Education.

Klapwijk, R. & Zhu, C. (2025). Spatial reasoning in explorative, creative, and playful design and science learning. In S. Earle, C. Preston, H. Georgiou, & A. Fitzgerald (Eds.), Primary science learning for children, teachers, and communities: Stories of practice and possibility for science educators. Springer Nature Singapore.

Zhu, C., Jolles, D., Klapwijk, R., Spandaw, J., de Vries, M. J. (under review). Spatial ability's contribution to pupils' creative design performance: Differential roles in ideation and prototyping.

Conference presentations

Buckley, J., & Zhu, C. (2025, September). Investigating the impact of comparison difficulty and judge fatigue on the feasibility of comparative judgement for technology education [Paper presentation]. The 42nd International Pupils' Attitudes Towards Technology (PATT) Research Conference (pp. 393-411). Montreal, Canada.

Zhu, C., & Klapwijk, R. (2024, October). Assessing pupils' creative design ideas and prototypes using comparative judgment [Paper presentation]. The 41st Pupils' Attitudes Towards Technology (PATT) Research Conference (pp. 290-296). Nanjing, China.

Zhu, C., & Klapwijk, R. (2024, June). The spatial aspect of designing: opportunities, challenges, and conjectures on engaging pupils in spatial thinking through design education [Paper presentation]. Spatial Cognition 2024 (pp. 97-113). Dublin, Ireland.

Zhu, C., & Klapwijk, R. (2023, June). Thinking spatially about data: a developing framework to understand children's spatial reasoning in data physicalization [Poster presentation]. The 22nd Annual ACM Interaction Design and Children (IDC) Conference (pp. 537-542). Chicago, USA.

Zhu, C., & Klapwijk, R. (2022, June). Scaffolding pupils' spatial thinking through design: A biomimicry project for the primary classroom [Paper presentation]. The 39th International Pupils' Attitudes Towards Technology (PATT) Research Conference (pp. 142-154). St. John's, Canada.