

Research Article

Magnetic Misconceptions in Kindergarten: Prior Conceptions, Conceptual Change, and the Role of Sociocultural Background

Despina Kozikopoulou ¹  and Konstantinos T. Kotsis ^{1,*} ¹ Lab of Physics Education and Teaching, Department of Primary Education, University of Ioannina, Greece* Correspondence: kkotsis@uoi.gr<https://doi.org/10.59652/azmrrb81>

Abstract: Research on early childhood science education has shown that young children often hold persistent alternative conceptions about magnetism, yet the relationship between sociocultural background and the qualitative structure of these misconceptions remains underexplored. This study investigates four magnetism misconceptions, animistic attribution, the all-metals belief, the permeability misconception, and the size-strength equation, among 50 Greek kindergarten children from two contrasting sociocultural groups: 25 Roma children from low sociocultural environments and 25 general-population children from middle-to-high sociocultural backgrounds. A qualitative-dominant mixed methods design was employed, combining individual diagnostic clinical interviews before and after instruction with field observations collected during a six-lesson Inquiry-Based Learning Teaching-Learning Sequence. Children's responses were analysed through a four-category reasoning scheme: absent framework, perceptual-associative reasoning, empirical reasoning, and protoscientific reasoning. Findings show that the two groups entered instruction with qualitatively different epistemic configurations. Following the intervention, surface accuracy improved substantially in both groups, but the depth of conceptual change differed: Roma children mainly achieved empirical revision based on observed experience, whereas general-population children more often developed protoscientific causal explanations. The size-strength misconception was universally held and revised through vivid disconfirmation, while the all-metals misconception remained the most resistant. The findings highlight the need for misconception-targeted, sequential, and socioculturally responsive early science instruction.

Keywords: conceptual change; magnetism misconceptions; preschool science education; Roma children; sociocultural background

Received: May 8, 2026

Accepted: July 25, 2026

Published: August 7, 2026



Copyright: © 2022 by the authors.
Submitted for open access publication
under the terms and conditions of the
Creative Commons Attribution (CC BY)
license
(<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The study of children's alternative conceptions in science has constituted one of the most productive research programmes in science education over the past five decades. Beginning with the foundational work of Driver and Easley (1978) and expanded by the conceptual change framework of Posner et al. (1982), the field has documented with remarkable consistency that children arrive at formal schooling not as blank slates but as holders of coherent, if scientifically inaccurate, theories about the physical world. These alternative conceptions, variously termed misconceptions, naïve theories, or prior conceptions in the literature, are not random errors but stable cognitive constructions that children have built through direct experience with the physical environment and through the analogical reasoning characteristic of preoperational thought (Piaget, 1989; Driver et al., 1994; Smolleck, 2011; Kotsis, 2023). Understanding the structure and persistence of such conceptions and designing instruction capable of dislodging them in favour of more scientifically adequate alternatives, remains one of the central challenges of early childhood science education.

Magnetism is among the most extensively studied domains for misconception research in young children, and for good reason. The phenomenon combines high intrinsic appeal, children are drawn to magnets from a very early age, with conceptual properties that are systematically misrepresented by naïve physical reasoning. The magnetic force operates at a

distance, through invisible fields, and with a selectivity that defies the most common intuitive shortcut children apply to it, namely the assumption that metallic appearance is the relevant criterion for magnetic susceptibility (Barrow, 2000; Bar et al., 1997; Ngazizah & Fatmaryanti, 2025). The result is a well-documented constellation of misconceptions that have been identified across cultures, age groups, and instructional contexts: the all-metals belief, the size-strength equation, the permeability misconception, and the animistic attribution of intentional agency to the magnet's pull (Ravanis, 1994; Kalogiannakis et al., 2018; Samara & Kotsis, 2024a). What is less well understood is the degree to which the structure of these misconceptions, their prevalence, their relative resistance to change, and their interaction with instructional treatment, is uniform across children of different sociocultural backgrounds or whether prior cultural exposure modulates the misconception profile in systematic and pedagogically relevant ways.

This gap in the literature is consequential. If magnetism misconceptions are truly universal in their distribution and structure, then the same instructional strategy should be equally effective for all children regardless of background. If, however, the misconception profile is shaped by the quality and type of prior informal experience with magnets, which is itself distributed unequally across sociocultural lines, then the pedagogical task is more differentiated than the existing literature has generally acknowledged. Some researchers (Morgan et al., 2016; Betancur et al., 2018) have demonstrated that socioeconomic disparities in science achievement are mediated primarily through differential access to informal science learning at home. Bourdieu's (1977, 1986) theory of cultural capital provides a theoretical framework for understanding how this differential access produces not merely quantitative differences in prior knowledge but qualitative differences in the cognitive schema with which children approach new scientific content. A child who has encountered magnets as refrigerator decorations, as toys, or in the context of parental explanation of physical principles arrives at the science classroom with a partially scaffolded conceptual structure. A child who has had no such exposure arrives with a conceptual vacuum that will be filled, in the absence of instruction, by animistic or purely associative reasoning. These two starting points are not equivalent, and the conceptual change process that instruction must initiate is correspondingly different in each case.

The present study investigates the nature and prevalence of magnetism misconceptions in two groups of kindergarten children distinguished by their sociocultural background, 25 Roma children from a low sociocultural environment and 25 children from the general population, and examines the differential processes of conceptual change that a structured Inquiry-Based Learning (IBL) Teaching-Learning Sequence (TLS) sets in motion across these groups. The specific misconceptions targeted were: the belief that magnets attract all metals regardless of composition (ALL-METALS); the belief that a physically larger magnet is necessarily stronger (SIZE-STRENGTH); the belief that an interposed medium such as water or plastic prevents magnetic field action (PERMEABILITY); and the attribution of the magnet's action to magical or animistic causes (ANIMISM). The study asks three interrelated questions. First, do Roma and general-population preschoolers differ in the prevalence and patterning of magnetism misconceptions at the point of school entry? Second, which specific misconceptions prove most resistant to conceptual change across both groups? Third, does the IBL TLS produce qualitatively different conceptual change trajectories in the two groups, and if so, what pedagogical and theoretical inferences can be drawn from those differences?

By attending to both the content and the process of conceptual change, rather than simply to pre- and post-test performance totals, this study contributes a more finely grained account of how sociocultural background interacts with science instruction in the preschool years. Its findings extend the growing body of literature on magnetism misconceptions in early childhood (Ravanis, 1994; Bar et al., 1997; Kalogiannakis et al., 2018; Samara & Kotsis, 2024a) by adding a sociocultural dimension that has hitherto been largely absent from that tradition, and they contribute to the broader theoretical discussion of conceptual change as a process that is not socioculturally neutral (Posner et al., 1982; Vosniadou, 1994; Chi, 2008).

Accordingly, the aim of the present study is to investigate how sociocultural background is associated with the prevalence, structure, and conceptual revision of young children's alternative conceptions about magnetism. More specifically, the study examines four recurring misconceptions – animistic attribution, the belief that magnets attract all metals, the permeability misconception, and the assumption that larger magnets are necessarily stronger – among Roma children from low-sociocultural environments and children from the general population with middle-to-high sociocultural backgrounds. It also explores how these conceptions develop following participation in a structured IBL TLS. The study is guided by

the following research questions:

RQ1: How do the two sociocultural groups differ in the prevalence and qualitative structure of their magnetism misconceptions before instruction?

RQ2: Which of the four targeted misconceptions are most resistant to conceptual change following the intervention?

RQ3: To what extent does the IBL TLS generate different qualitative conceptual change trajectories in the two groups?

2. Theoretical Framework

2.1. Conceptual Change as Theoretical Foundations and the Status of Misconceptions

The concept of conceptual change has generated one of the richest and most contested theoretical traditions in science education research. At its foundation lies the recognition, articulated most influentially by Posner et al. (1982) and Liljedahl (2011), that meaningful science learning is not a matter of adding new information to an empty cognitive store but of revising, restructuring, or replacing prior knowledge that is already there, since alternative ideas can function as productive starting points for physics teaching when they are explicitly identified and pedagogically addressed. Posner et al. (1982) drew explicitly on Kuhnian philosophy of science to argue that the learner, like the scientific community, operates with an existing conceptual framework within which anomalous data must either be assimilated or, when the conditions for conceptual change are met, used as the catalyst for a genuine accommodation of the kind Piaget (1952) described. For conceptual change to occur, Posner et al. (1982) proposed four conditions. They state the following: the learner must experience dissatisfaction with the current conception; an intelligible, plausible, and fruitful alternative must be available; and the alternative must appear capable of resolving the anomaly. This framework has proved enormously generative for the design of science instruction, because it foregrounds the teacher's task as the engineering of productive dissatisfaction, the creation of conditions in which children's existing conceptions are brought into direct confrontation with evidence they cannot easily accommodate.

Within this broad tradition, two theoretical positions have dominated subsequent debate and carry particular relevance for the present study. The first is Vosniadou's (1994, 2013) framework theory approach, which proposes that children's prior conceptions are not isolated beliefs but elements of coherent, if implicit, framework theories that function as explanatory systems for understanding entire domains of physical experience. On this account, misconceptions in science are not random errors but predictable products of a naive framework theory that is internally consistent and resistant to local correction precisely because any revision of one element tends to destabilise the broader framework within which it is embedded. The all-metals misconception in magnetism, for example, is not simply an incorrect factual belief; it is an expression of a deeper framework assumption that perceptual surface features, particularly metallic appearance, reliably predict physical behaviour. Instructional strategies that target only the surface belief without addressing the underlying framework will tend to produce fragile or unstable learning, because the corrected belief lacks the conceptual scaffolding that would make it durable (Vosniadou, 2013). The implication for pedagogical design is that effective conceptual change instruction must make the framework assumption itself visible and subject to revision, typically through the engineering of multiple, varied instances of cognitive conflict that collectively challenge the broader inferential pattern rather than any single belief.

The second major position is that of Chi (2008), who distinguishes three qualitatively different types of conceptual change, belief revision, mental model transformation, and categorical shift, and argues that the type of change required depends on the ontological status of the prior conception.

Belief revision, the least demanding type, involves correcting a factual error within an existing conceptual category; it requires no fundamental reorganisation of knowledge and is achievable through direct counter-evidence. Mental model transformation requires the restructuring of an entire explanatory schema, such as the shift from a contact-force model of magnetic action to an action-at-a-distance model. Categorical shift, the most demanding type, requires moving a concept from one ontological category to another, for instance from the category of substances or objects, where children tend to place forces, to the category of processes or fields. Chi's taxonomy is useful for the present study because it predicts the relative ease and resistance to change of different magnetism misconceptions. The all-metals belief and the size-strength equation are candidates for belief revision or, at most, mental

model transformation, and should therefore be addressable through hands-on disconfirmation. The persistence of size-based reasoning and the broader difficulty of constructing coherent mental models of magnetism have also been documented with older students, where magnetism has been examined both as a size-dependent property within learning progressions from familiar magnetic phenomena toward nanoscale science and as a domain in which students develop varied explanatory models of magnetic interaction (Naushabekov et al., 2025; Kähkönen et al., 2020; Varela-Aldás et al., 2026). The attribution of magnetic action to magical or animistic agency, by contrast, may require a more profound categorical shift, from intentional entities to physical processes, that cannot be accomplished within a single intervention and that may produce, as this study's data suggest, a distinctive intermediate state in which children adopt empirically accurate responses while retaining animistic causal language.

A complementary perspective is provided by the knowledge-in-pieces framework, which offers an alternative account of the structure of prior physical knowledge (diSessa, 1993). Rather than coherent framework theories, it also proposes that intuitive physics consists of a large repertoire of phenomenological primitives, or p-prims, which are minimal, context-sensitive knowledge structures abstracted from direct physical experience. P-prims are neither correct nor incorrect in themselves; they are applied appropriately or inappropriately depending on the situational cues that activate them. On this account, the size-strength misconception in magnetism is not a coherent belief but the activation of a highly general and normally productive p-prim, "bigger is more powerful", in a context where it does not apply. The p-prim itself is not wrong; it correctly describes the relationship between size and force in many familiar physical contexts (a bigger hammer hits harder; a larger engine produces more power). Its misapplication to magnets is a matter of inappropriate contextual activation, not systematic theoretical error. This interpretation is consistent with the view that learners' explanatory models often integrate both conscious and intuitive conceptual resources, making it pedagogically important to identify, refine, and reorganise such resources rather than simply replace them as errors (Cheng & Brown, 2010; Graziano, 2022). The instructional implication is correspondingly different from that of Vosniadou's (1994) framework; rather than dismantling a theory, the teacher's task is to help children discriminate the contexts in which the p-prim applies from those in which it does not, by creating situational contrast experiences that make the contextual boundaries salient.

These three theoretical frameworks, the framework theory approach (Vosniadou, 1994, 2013), the ontological categorisation model (Chi, 2008), and the knowledge-in-pieces approach (diSessa, 1993), are not mutually exclusive. They capture different grain sizes of the same cognitive territory and, taken together, provide a richer account of the conceptual change task than any one of them offers alone. Their integration informs the analytical framework of this study, which examines not only whether children's responses changed between pre- and post-test but the qualitative nature of that change in case it constitutes belief revision, partial framework accommodation, or the more limited achievement of correct behaviour without commensurate categorical reorganisation.

2.2. *The Specific Misconception Landscape of Magnetism in Early Childhood*

Against this theoretical backdrop, the specific misconception landscape of magnetism in preschool-age children has been mapped with considerable precision. Ravanis (1994), in the foundational empirical study of this domain within a Piagetian framework, documented that five-year-old children already hold stable, if scientifically inaccurate, ideas about the selective properties of magnets, and that free exploratory activity with magnets in a material-rich environment is insufficient on its own to produce conceptual revision without structured mediation. Similarly, Kakana et al. (2009) as well as McCormick et al. (2000) showed that teaching magnetic attraction to preschool children depends on the instructional approach adopted, further supporting the need for carefully designed and mediated teaching interventions.

Subsequent work by Bar et al. (1997) identified the permeability misconception, the belief that interposed media such as water or solid objects block magnetic action, as one of the most cross-culturally stable errors in children's understanding of action at a distance, attributable to the general p-prim that physical barriers prevent physical action. Driver et al. (1994) provided the broader catalogue of alternative conceptions in secondary-level science, within which the all-metals and size-strength misconceptions for magnetism represent some of the most deeply rooted and pedagogically challenging examples. Kalogiannakis et al. (2018) confirmed the persistence and distribution of these misconceptions in contemporary Greek

preschool classrooms, finding that children attributed magnetic attraction to the magnet's miraculous or magical qualities, an animistic attribution consistent with Piaget's (1989) characterisation of preoperational reasoning, even after exposure to narrative-based instruction. Samara and Kotsis (2023) demonstrated that STEM-integrated hands-on interventions can produce significant post-test gains in kindergarten children's understanding of magnetism, although the all-metals classification remained one of the most resistant targets, with some children continuing to include non-ferrous metals in the magnetic category after instruction. This finding is consistent with Chi's (2008) framework. The all-metals belief is embedded in a broader perceptual schema in which metallic surface appearance carries reliable predictive weight, making it more resistant to revision than the size-strength equation, which can be directly and dramatically disconfirmed through a single comparative experiment.

2.3. Sociocultural Background as a Moderator of the Misconception Profile

What the existing magnetism misconception literature has not systematically addressed is the question of whether sociocultural background modulates the structure or prevalence of these prior conceptions. The assumption, implicit in most of this literature, is that magnetism misconceptions are developmentally universal, products of Piagetian preoperational reasoning or diSessa's p-prims that emerge independently of cultural context. This assumption is only partially warranted. While the cognitive mechanisms underlying certain misconceptions, such as the size-strength equation or animistic attribution, appear to be genuinely universal, the distribution of other misconceptions is sensitive to prior informal experience, which is itself socioculturally structured. A child who has played with magnets at home, observed adults using magnetic tools, or received adult-mediated explanation of magnetic selectivity arrives at school with a partial conceptual scaffold that a child without any such exposure simply does not possess. This is not a matter of cognitive readiness in the Piagetian sense; it is a matter of differential access to the informal science learning environments through which the beginnings of a conceptually accurate framework are assembled (Betancur et al., 2018; Bourdieu, 1986; Cabra Hernández, 2025).

The consequence for misconception research is that two children who hold the all-metals misconception at school entry may be in qualitatively different epistemic positions. The child from a high-capital home environment holds the misconception as an overgeneralisation from prior partial knowledge; she has encountered magnets, knows they do something, and has incorrectly extended what she knows to all metals. The child from a low-capital environment, by contrast, may hold what is better characterised as an absence of any framework rather than a misconception in the technical sense. Her prior cognitive schema for the domain contains no organised structure to be revised, only the general-purpose reasoning heuristics of preoperational thought. For the former, conceptual change instruction must produce a restriction and refinement of an existing category. For the latter, it must build an appropriate category from the ground up while simultaneously constraining the general-purpose heuristics that the absence of domain-specific knowledge leaves in play. These are pedagogically distinct tasks, even when they appear superficially similar on a pre-test instrument, and their differential character has significant implications for both the design of IBL interventions and the interpretation of their outcomes.

2.4. IBL as the Instructional Vehicle for Conceptual Change

The selection of IBL as the instructional vehicle for the present study is grounded in the conceptual change literature as well as in the broader science education research reviewed in the companion paper. IBL creates the conditions for productive conceptual conflict, prediction, experimentation, observation, and guided reflection, that Posner et al. (1982) identified as necessary for accommodation, while experiments in inquiry-based science teaching provide learners with direct empirical experiences through which prior conceptions can be tested, challenged, and reconstructed (Kotsis, 2024). The hands-on structure of IBL activities means that the empirical disconfirmation of a misconception is produced by the child's own investigative action rather than by teacher testimony, a feature that is theoretically significant because it addresses the intelligibility and plausibility requirements of the Posner et al.'s (1982) framework stating that a child who has personally produced the disconfirming evidence cannot easily dismiss it as incomprehensible or implausible. In this direction, the use of new technologies, robotics, and STEM-oriented activities in magnetism teaching can further enrich children's opportunities for active exploration, classification, and conceptual reconstruction (Samara & Kotsis, 2023).

Furthermore, the cycle of prediction followed by observation is a direct pedagogical

implementation of the diSession insight that contextual contrast, experiencing the same p-prim as productive in some contexts and failing in others, is more effective for adjusting the activation conditions of p-prims than is direct correction of stated beliefs (diSessa, 1993). For children from low sociocultural backgrounds who lack prior domain-specific knowledge, IBL's additional advantage is that its entry point is direct physical experience rather than prior linguistic or declarative knowledge, thereby lowering the effective threshold of cultural capital required for participation (French, 2004; Ravanis, 2022; Sowmya & Rani, 2025). In this context, artificial-intelligence-supported instructional design can also assist early childhood educators in planning magnetism activities that are developmentally appropriate, inquiry-oriented, and responsive to children's prior conceptions (Samara & Kotsis, 2024b).

Taken together, these theoretical perspectives suggest that conceptual change in early magnetism education cannot be understood solely as a shift from incorrect to correct answers. It must also be examined as a qualitative transformation in the reasoning structures through which children interpret magnetic phenomena. The nature of this transformation may depend both on the type of misconception being addressed and on the sociocultural experiences that have shaped children's prior conceptual resources. Within this framework, IBL provides an appropriate instructional context for making prior ideas visible, confronting them with observable evidence, and supporting progressively more scientific explanations. These assumptions guide the methodological design and the analysis of children's responses before, during, and after the intervention.

3. Materials and Methods

This study employs a qualitative-dominant mixed methods design with a misconception-diagnostic focus. While it draws on the same empirical dataset as the companion paper reporting aggregate pre- and post-test outcomes across two sociocultural groups, its analytical orientation is fundamentally different. Rather than asking whether the TLS reduced the performance gap between groups, it asks what specific misconceptions children held, how those misconceptions were structured and distributed across groups, and what the qualitative character of the conceptual change process was for each group following the intervention. This distinction in analytical purpose necessitates a corresponding shift in methodology, from the inferential statistical analysis appropriate to comparative outcome research toward the qualitative analysis of children's reasoning, justification, and causal language as they evolve through the pre-test, intervention, and post-test phases.

The methodological tradition informing this approach is the Piagetian clinical interview, which has served as the foundational instrument for research on children's prior conceptions in science since the earliest systematic investigations of children's physical reasoning (Piaget, 1989). The clinical interview method prioritises the structure and quality of children's thinking over the mere correctness of their answers, attending specifically to the justifications children offer for their responses, the causal vocabulary they employ, and the consistency of their reasoning across related questions (Ravanis, 1994; Ginsburg, 1997, as cited in argumentation research). This approach is particularly well suited to misconception research because, as Vosniadou (1994) has argued, the surface-level correctness or incorrectness of a response is often a poor guide to the underlying conceptual structure it expresses. A child may produce a correct prediction for incorrect reasons, or may hold a partially revised conception that correct-only scoring would either credit or penalise depending on the specific question wording. Attending to the reasoning behind responses, rather than only to the responses themselves, is therefore not merely preferable but methodologically necessary for the kind of analysis this study undertakes.

The participant sample was identical to that described in the companion study that 50 pre-school children aged four to six years, drawn from two public kindergartens in an urban area of Greece. Group A comprised 25 Roma and Aromunian children from a school characterised by a uniformly low sociocultural profile; Group B comprised 25 children from the general population with middle-to-high sociocultural backgrounds.

Purposive sampling was employed on the basis of the theoretically motivated contrast between the two groups' sociocultural profiles, following established practice in qualitative and mixed-methods educational research (Creswell, 2013). All children participated in the same six-lesson IBL TLS as described in the companion paper. Ethical approval and parental consent were obtained in accordance with institutional requirements prior to the commencement of data collection.

Three instruments generated the data analysed in this study. The primary instrument was

the individual semi-structured diagnostic interview, administered at two time points: immediately before the intervention (pre-test) and immediately after its completion (post-test). The interview was conducted by the researcher-teacher individually with each child in a quiet space adjacent to the classroom, using physical objects rather than verbal descriptions as the primary stimulus materials. This design choice was deliberate and methodologically significant for children at the preoperational stage of development, and particularly for children from linguistically marginalised backgrounds, abstract verbal questioning about physical phenomena invites confabulation or acquiescence rather than genuine expression of prior knowledge. The use of physical objects, actual magnets, collections of metal and non-metal items, a transparent container of water with submerged paper clips, magnets of varying sizes and verified strengths, grounded each question in direct perceptual experience and thereby minimised the confounding of linguistic competence with conceptual understanding (DeCosta et al., 2023; Ponizovsky-Bergelson et al., 2019).

The interview addressed four conceptual targets, each of which mapped directly onto a specific misconception identified in the literature. The first target, magnet recognition and animistic attribution (ANIMISM / RECOGNITION), was assessed through open-ended elicitation. The child was shown the magnet and asked what it was, what it could do, and why it did what it did. This question was designed not merely to establish whether the child recognised the object but to surface the causal framework, scientific, magical, or absent, within which the child located its action. The second target, material classification and the all-metals misconception (ALL-METALS), was assessed by presenting eight objects, a steel paper clip, an iron nail, an aluminium strip, a plastic button, a wooden block, a copper coin, a glass marble, and a steel bolt, and asking the child to predict which ones the magnet would attract, before and after testing each object directly. The pre-test prediction data and the post-test prediction data were both recorded, along with any spontaneous justification the child offered. The third target, the permeability misconception (PERMEABILITY), was assessed by showing the child a paper clip submerged in a transparent container of water and asking whether the magnet held outside the container would be able to attract it, along with an explanation of the predicted outcome. The fourth target, the size-strength equation (SIZE-STRENGTH), was assessed by presenting two magnets of markedly different physical sizes but pre-verified, deliberately reversed strengths, the smaller magnet being measurably stronger, and asking the child which would attract more paper clips and why.

All responses, including verbatim causal justifications, were recorded contemporaneously by the researcher using a structured response protocol sheet. Responses were subsequently coded along two dimensions: accuracy (correct / incorrect) and reasoning type. The reasoning type dimension was the primary analytical focus of the present study and distinguished among the following categories derived inductively from the data and deductively from the conceptual change literature: animistic or magical reasoning (attributing the magnet's action to supernatural or intentional agency); perceptual-associative reasoning (attributing predictions to surface perceptual features such as metallic appearance, colour, or size); empirical reasoning (attributing outcomes to directly observed experimental results without causal explanation); and pro-to-scientific reasoning (offering a causal explanation that approximates or correctly invokes physical principles such as magnetic force, attraction, or field). These four reasoning categories correspond approximately to the progression from naïve framework theory through partial revision to genuine conceptual change as described by Vosniadou (1994, 2013), and to Chi's (2008) three-level taxonomy of belief revision, mental model transformation, and categorical shift.

The second instrument was the researcher's systematic field observation journal, maintained throughout the six intervention lessons. Entries recorded children's spontaneous verbal responses during the activities, instances of expressed surprise or cognitive conflict, modifications of initial predictions in real time, peer interaction around contested observations, and linguistic markers of emerging conceptual revision. This journal provided the process-level data that the pre-test/post-test interview structure could not capture. The moment-by-moment trajectory of conceptual negotiation during the intervention itself, which is invisible to any assessment instrument administered before or after instruction.

The third instrument was the collection of children's illustrated non-verbal worksheets completed during the intervention lessons. Where verbal reasoning was difficult to elicit, particularly for Group A children in the early lessons, worksheet responses provided supplementary evidence of children's classification behaviour and their evolving understanding of magnetic selectivity.

Data analysis proceeded in three phases. In the first phase, pre-test interview responses

for all 50 children were systematically coded for accuracy and reasoning type on each of the four conceptual targets independently by the researcher. Coded responses were then compared across the two groups to establish the pre-intervention misconception profile of each group; not merely how many children held each misconception, but what reasoning type underpinned it. This phase addressed the first research question, concerning whether Roma and general-population preschoolers differed in the prevalence and patterning of misconceptions at school entry.

In the second phase, post-test responses were coded using the same scheme and compared with the corresponding pre-test responses for each child. This individual-level comparison enabled the identification of three distinct change patterns: full conceptual revision (from incorrect reasoning to proto-scientific reasoning), partial revision (from incorrect reasoning to empirical reasoning, or from a more naïve to a less naïve incorrect reasoning type), and non-revision (no change in reasoning type despite a change in surface response accuracy). The distribution of these change patterns across the two groups addressed the second and third research questions, concerning the relative resistance of different misconceptions to change and the qualitative differences in conceptual change trajectories between groups.

In the third phase, field observation journal entries and worksheet data were integrated with the interview data to produce a narrative account of the conceptual change process for each misconception within each group. This integration allowed the analysis to distinguish, for example, between children whose post-test accuracy reflected genuine revision of the underlying causal schema and children whose correct post-test response was anchored solely in episodic memory of an observed experimental outcome, a distinction that has significant implications for the durability of learning but that surface-level accuracy data cannot reveal.

Several methodological limitations of the present analytical approach require acknowledgement. The coding of reasoning types from verbatim transcripts of preschool-age children's responses inevitably involves interpretive judgement, and the absence of a second independent coder for reliability verification is a limitation of the study's qualitative claims. The four-category reasoning scheme, while derived from both the empirical data and established theoretical frameworks, is necessarily a simplification of a more continuous and variable cognitive process; children's reasoning in practice frequently combined elements of more than one category within a single response, and such hybrid responses required allocation to the dominant reasoning type. Furthermore, the post-test was administered immediately following the intervention, meaning that the reasoning types observed at post-test reflect the cognitive state at the moment of instruction's conclusion; they do not guarantee stable long-term revision of the underlying conceptual framework, a point that the conceptual change literature emphasises as one of the most persistent challenges in science education research (Vosniadou, 2013; Chi, 2008).

4. Results

This section presents the findings of the misconception analysis in three parts: the pre-test misconception profile, describing the nature and reasoning type underlying each prior conception in both groups; the post-test change typology, characterising the qualitative nature of conceptual revision achieved following the TLS; and an integrated account of misconception resistance and the differential change trajectories of the two groups. All data are drawn from the individual diagnostic interview administered before and after the intervention, supplemented by field observation notes and worksheet analysis.

4.1. Pre-Test Misconception Profiles

Table 1 presents the pre-test misconception profile of both groups across all four conceptual targets, organised by misconception and dominant reasoning type. The pre-test data reveal that the two groups entered the intervention with not only different levels of accuracy but qualitatively different misconception structures. For the animistic attribution target, 23 of 25 Group A children (92%) operated without any prior domain-specific framework. It refers to the fact that they had never encountered the word “magnet”, had no stored conceptual schema for its action, and attributed its observed effects to the teacher's supernatural agency. The representative response, “The teacher is doing magic”, is not a misconception in the strict theoretical sense of an incorrect belief held within an existing framework; it is the output of preoperational animistic reasoning applied to a completely novel phenomenon in the absence of any prior knowledge that could scaffold a more

differentiated response. This is closer to what Chi (2008) would characterise as a categorical gap, the complete absence of the relevant ontological category, than to the belief revision scenario that most misconception research addresses. Group B, by contrast, was uniformly correct on this target at pre-test, with all 25 children identifying the magnet by name and attributing its action, at least partially, to a physical force.

Table 1. Pre-test misconception prevalence and dominant reasoning type by group (n = 50).

Misconception	Pre-test reasoning type	Group A – Roma (n = 25)	Group B – General (n = 25)
Magnet recognition / Animistic attribution	Absent framework: magical / animistic	23/25 (92%)	0/25 (0%)
	Perceptual-associative (object recognised, partial knowledge)	2/25 (8%)	25/25 (100%)
	Perceptual-associative (metallic appearance as criterion)	24/25 (96%)	3/25 (12%)
All-metals misconception (material classification)	Proto-scientific (ferrous/non-ferrous distinction)	0/25 (0%)	22/25 (88%)
	Absent framework (no prior contact with magnets)	1/25 (4%)	0/25 (0%)
Permeability misconception (magnetic field through water/plastic)	Perceptual-associative (barrier blocks all physical action)	8/25 (32%)	4/25 (16%)
	Proto-scientific (force acts through media)	17/25 (68%)	21/25 (84%)
Size-strength equation (larger magnet = stronger)	Perceptual-associative (size as proxy for power, universal p-prim)	25/25 (100%)	25/25 (100%)

Note: Reasoning type categories: Absent framework = no prior domain-specific knowledge; ani-mistic or magical causal attribution. Perceptual-associative = incorrect prediction based on surface perceptual features (metallic appearance, object size). Proto-scientific = reasoning that approximates correct causal mechanism. Categories are mutually exclusive per child per misconception.

For material classification, the all-metals misconception, the structural difference between the groups is equally significant but differently configured. In Group A, 24 of 25 children (96%) held an essentially perceptual-associative schema – they predicted that all shiny, hard objects would be attracted to the magnet, without discrimination between ferrous and non-ferrous metals. The single child who performed correctly did so on the basis of prior direct experience rather than conceptual understanding. In Group B, 22 of 25 children (88%) correctly classified the ferrous and non-ferrous materials before any instruction, demonstrating that their prior informal exposure had produced a functionally accurate framework for material selectivity. The three Group B children who held the all-metals misconception did so in the specific form of including aluminium foil among attracted materials, a perceptual-associative error driven by the high-salience shininess of foil rather than a global failure of material discrimination. This is a qualitatively different, and substantially less entrenched, version of the misconception than that held by Group A, being a local overgeneralisation within an otherwise accurate partial framework rather than a wholesale reliance on perceptual surface features.

The permeability misconception showed the most complex and, in some respects, counter-intuitive pre-test pattern. Group A showed higher rates of the correct conception (68% predicted the magnetic field would operate through water) than might be expected given their generally lower prior knowledge. This apparent paradox is explicable within the diSession knowledge-in-pieces framework. In particular, the belief that forces operate through physical barriers is not peculiarly associated with science education background but reflects the activation of a relatively common everyday p-prim, forces act at a distance and penetrate obstacles, that is applied appropriately here by children who have no specific misconception schema for magnetic fields to override it. The children who held the incorrect conception (32% of Group A, 16% of Group B) were activating a competing and, in most everyday physical contexts, more salient p-prim that solid or liquid barriers prevent physical action. Notably, this same competing p-prim was held by the majority of adult educators present during the intervention, four out of six, confirming that the permeability misconception is not a product of naïve physical reasoning alone but is widely distributed across the general adult population.

The size-strength equation stood apart from all other misconceptions in one decisive respect. It was held universally, by all 50 children in both groups, with identical reasoning type. Every child, regardless of sociocultural background, prior knowledge level, or any other individual characteristic, predicted that the physically larger magnet would attract more paper clips because it was bigger. This universality is the empirical signature of a phenomenological primitive in diSessa's (1993) sense, particularly the "bigger is more powerful" p-prim is so deeply embedded in children's physical intuitions, and so broadly confirmed by everyday experience, that no amount of informal prior knowledge about magnets specifically is sufficient to override it. Its eradication through a single comparative experiment was correspondingly dramatic in both groups, producing visible surprise and immediate revision, the clearest instance of productive cognitive conflict in the entire dataset.

4.2. Post-Test Reasoning Type Distribution

Table 2 presents the post-test reasoning type distribution following the TLS, using the same four-category scheme applied at pre-test.

Table 2. Post-test reasoning type distribution by misconception and group.

Misconception	Post-test reasoning type	Group A – Roma (n = 25)	Group B – General (n = 25)
Magnet recognition / Animistic attribution	Proto-scientific (names magnet, cites attraction)	25/25 (100%)	25/25 (100%)
	Empirical ("I saw it stick", correct, no causal schema)	Dominant in Group A	Minority in Group B
All-metals misconception (material classification)	Proto-scientific (correct classification with justification)	13/25 (52%) [full attenders only]	25/25 (100%)
	Perceptual-associative (partial revision, ferrous metals correct, non-ferrous still uncertain)	12/25 (48%) [irregular attenders]	0/25 (0%)
Permeability misconception (magnetic field through water/plastic)	Proto-scientific (force passes through media)	25/25 (100%)	25/25 (100%)
Size–strength equation (larger magnet = stronger)	Empirical ("I saw the small one hold more", correct, but p-prim not explicitly rejected)	25/25 (100%)	25/25 (100%)
	Proto-scientific (explicit rejection of size-power link)	Minority in Group A	Majority in Group B

Note: Full attenders = Group A children who attended all six intervention lessons without absence (n = 13). Irregular attenders = Group A children who missed one or more sessions (n = 12). Group B had no attendance irregularities across the intervention week.

The most theoretically significant pattern in the post-test data is the consistent divergence between the two groups in the qualitative character of their correct responses, which is most clearly visible in the animistic attribution and size-strength targets. On the animistic attribution target, all 25 Group A children achieved correct recognition of the magnet at post-test, a complete reversal of their pre-test profile, but the causal reasoning underlying this correct response remained predominantly empirical rather than proto-scientific. When asked why the magnet attracted certain objects, the dominant Group A response was experiential and non-explanatory: "Well, I saw it stick", "It pulls, I know because it did it". These responses are correct, and they represent a genuine abandonment of the animistic attribution that dominated at pre-test. However, they do not yet involve the construction of a causal schema that could generalise beyond the immediately experienced cases. In Chi's (2008) terms, belief revision has been achieved, the incorrect belief has been replaced, but the categorical reconstitution that would make the new belief durable has not. Group B children, by contrast, predominantly produced proto-scientific responses at post-test, citing "force", "attraction", or "magnetic power" as the causal mechanism, demonstrating that their prior partial framework provided the scaffolding necessary to integrate the intervention's experimental findings into a generalisable causal account.

On material classification, the post-test data reveal the attendance effect with particular clarity. The 13 Group A children who attended all six sessions of the intervention achieved correct classification of all eight materials, with justifications that, while predominantly empirical ("The nail yes, the foil no, I tried it"), indicated functional revision of the all-metals

schema. The 12 Group A children whose attendance was interrupted, missing at least one session, typically Lesson 2 or 3, which targeted material classification and the size-strength experiment respectively, failed to achieve correct classification, and their post-test responses continued to show perceptual-associative reasoning. This within-group pattern is the most direct evidence in the dataset that the failure to revise the all-metals misconception in a subset of Group A children was a product of sequential learning disruption rather than conceptual incapacity. Group B achieved universal correct classification with proto-scientific justification, consistent with their pre-test partial framework providing a foundation onto which the intervention's experimental evidence could be mapped.

The permeability target showed full convergence at post-test, with all 50 children across both groups achieving correct responses. The nature of the change, however, differed between groups in the manner already observed: Group A children relied predominantly on empirical justification ("I saw it pull the paper clip through the water"), while Group B children more frequently offered proto-scientific accounts ("The force goes through the plastic, it doesn't stop it"). The size-strength target also achieved universal post-test accuracy, but again with a qualitative divergence – the dominant Group A justification was pure empirical surprise ("The small one held more! I can't believe it"), while Group B children more frequently produced responses that approximated an explicit rejection of the size-power assumption ("Size doesn't make it stronger, it's the power inside that matters").

4.3. Conceptual Change Typology and Theoretical Interpretation

Table 3 synthesises the change patterns across all four misconceptions, characterising each in terms of the conceptual change typology developed in the theoretical framework and relating each pattern to the relevant theoretical account.

Table 3. Conceptual change typology by misconception and group, with theoretical interpretation.

Misconception	Group A change type	Group B change type	Theoretical interpretation
Animistic attribution (RECOGNITION)	Belief revision → empirical	None required (correct pre-test)	Belief revision achieved; animistic causal language abandoned but causal mechanism not fully reconstructed in Group A (Chi, 2008)
All-metals misconception (MATERIALS)	Partial revision (attenders) / Non-revision (absentees)	Full belief revision → proto-scientific	Framework assumption (metallic appearance = magnetic) requires sustained exposure for revision; single-session exposure insufficient for irregular attenders (Vosniadou, 1994; 2013)
Permeability misconception (WATER)	Belief revision → proto-scientific	Consolidation (majority already correct pre-test)	Barrier p-prim overridden by direct experimental observation; single vivid disconfirmation sufficient, lowest resistance misconception (diSessa, 1993)
Size-strength equation (POWER)	Belief revision → empirical (both groups)	Belief revision → proto-scientific (majority)	Universal p-prim; dramatic single disconfirmation produces correct behaviour in all children; Group B achieves partial categorical reframing, Group A anchors in episodic memory (diSessa, 1993; Chi, 2008):

Note: Change type labels follow Chi (2008), where belief revision = replacement of an incorrect belief within an existing category. Partial revision = movement from a more to a less naïve incorrect reasoning type without full correct conceptualisation. Non-revision = no change in reasoning type despite correct surface response (episodic memory only). Categorical reframing = partial restructuring of the ontological category within which the concept is held.

The typology in table 3 foregrounds a pattern that raw accuracy data would conceal. It means that the qualitative depth of conceptual change achieved by the two groups diverges systematically, and the direction of that divergence is consistent with the theoretical predictions of both Vosniadou's (1994; 2013) framework theory and Chi's (2008) ontological categorisation model. Where Group A children achieve belief revision, it tends to be anchored in episodic empirical memory rather than reconstructed causal schema, a fragile form of revision that is vulnerable to regression precisely because, as Vosniadou (2013) argues, beliefs that are not embedded in a revised framework lack the conceptual support that makes them

lasting. Group B children, whose pre-existing partial framework provided the scaffolding for integrating new experimental evidence, more frequently achieve what Chi (2008) would characterise as the beginning of categorical reframing, particularly on the size-strength and permeability targets, producing proto-scientific responses that suggest a more stable revision of the underlying conceptual structure.

4.4. Verbatim Reasoning Exemplars

Table 4 presents selected verbatim or near-verbatim reasoning exemplars drawn from the diagnostic interviews and field observation notes, illustrating the qualitative character of the reasoning types identified in tables 1 through 3.

Table 4. Selected verbatim reasoning exemplars by misconception, time point, and group.

Misconception	Time point	Group A – Roma (exemplar)	Group B – General population (exemplar)
Animistic attribution	Pre-test	“The teacher is doing magic”	“It’s a magnet, it pulls metal things”
	Post-test	“It sticks, I saw it” (empirical; causal schema absent)	“It has a force that attracts iron” (protoscientific)
All-metals misconception	Pre-test	Predicted: aluminium foil, copper coin, plastic, iron nail, all attract	Predicted: aluminium foil attracts (shiny metal heuristic); iron nail correct
	Post-test	(Consistent attenders): “The nail and the bolt yes, the foil and the coin no”	“Only the iron things, not all metals. The aluminium doesn’t”
Permeability misconception	Pre-test	“Maybe it will” (uncertain; 8 predicted it would not work)	“The force goes through, it will pull it” (majority correct pre-test)
	Post-test	“Yes, I saw it, it still pulled” (empirical)	“The force passes through the plastic and grabs the paper clip” (protoscientific)
Size-strength equation	Pre-test	“The big one, because it’s bigger it will pull more” (universal)	“The big one must be stronger” (universal)
	Post-test	“The small one! I saw it hold more” (empirical; surprise expressed)	“Size doesn’t matter, the small one was stronger in the experiment” (partial categorical reframing)

Note: All exemplars are drawn from verbatim or near-verbatim records in the researcher’s field observation journal and structured response protocol sheets. Group A exemplars are translated from Greek; original language preserved in the thesis archive. Exemplars were selected to represent the dominant reasoning type in each cell; atypical responses are discussed in the prose analysis.

The exemplars in table 4 illustrate with particular clarity the distinction between empirical and protoscientific reasoning that the quantitative change typology can only classify but not convey. The Group A post-test response to the animistic attribution target, “It sticks, I saw it”, and the Group B response, “It has a force that attracts iron”, are both correct. They are not, however, epistemically equivalent. The former records an observation; the latter constructs an explanation. Similarly, the Group A post-test response to the size-strength target, “The small one! I saw it hold more”, and the Group B partial reframing, “Size doesn’t matter, the small one was stronger in the experiment”, both correctly reject the pre-test prediction, but the Group B response moves toward explicit propositional revision of the size-power p-prim, while the Group A response leaves the p-prim intact and simply records its disconfirmation in this specific case. These subtle but consequential distinctions in the depth of conceptual change are precisely what the reasoning-type analytical framework was designed to capture, and they constitute the primary empirical contribution of this article to the magnetism misconception literature.

5. Discussion

The pre-test data presented in this study make clear that the concept of “misconception” is not a unitary category, and that treating it as such produces a systematic misreading of the pedagogical challenge faced in socially heterogeneous preschool classrooms. Among the four conceptual targets investigated, three distinct epistemic configurations were observed at pre-test, each of which corresponds to a different starting condition for conceptual change

instruction and requires a correspondingly different response.

The first configuration was the complete absence of a domain-specific framework, observed predominantly in Group A on the magnet recognition and material classification targets. The 23 Group A children who attributed the magnet's action to supernatural agency were not, in any precise theoretical sense, holding a misconception about magnetism. It shows that they had no prior conceptual structure for the domain within which a misconception could be lodged. They were operating with the general-purpose animistic and perceptual-associative reasoning of pre-operational thought, applied in the absence of any domain-specific knowledge that could constrain or scaffold a more differentiated response. This configuration is more accurately described, within Chi's (2008) framework, as an ontological gap, the complete absence of the relevant category, than as a belief error within an existing category. Its pedagogical implication is that instruction must accomplish something more fundamental than misconception remediation. It must build the category itself before any revision of beliefs within it can take place. That the TLS achieved this for all Group A children who attended consistently, producing universal correct recognition at post-test, is a strong demonstration of IBL capacity to construct domain-specific knowledge from the ground up, through direct physical experience, in children for whom no prior scaffolding exists.

The second configuration was the overgeneralisation of a partially correct framework, observed in Group B on the material classification target. The three Group B children who included aluminium foil among magnetically attracted materials were not operating without a framework; they were operating with a framework that was largely correct but locally over-extended, applying a high-salience perceptual heuristic, shininess as a marker of metallic, and therefore magnetic, character, in a case where it did not apply. This is the form of misconception that Vosniadou (1994, 2013) most precisely describes as a framework theory error. Particularly, the child has a genuine prior conceptual structure, the revision required is local rather than global, and the primary pedagogical task is to restrict the overgeneralisation by providing discriminating evidence. The TLS accomplished this reliably for all Group B children, consistent with the pre-diction that local framework errors are more tractable than categorical gaps.

The third configuration was the universal phenomenological primitive, exemplified by the size-strength equation. As the pre-test data confirm, the belief that a physically larger magnet is necessarily stronger was held by every child in both groups, regardless of sociocultural background, prior knowledge level, or familiarity with magnets. This universality is the defining signature of what diSessa (1993) termed a phenomenological primitive – a minimal, context-sensitive knowledge element abstracted from ubiquitous physical experience, in this case, the experience that larger objects routinely exert greater physical effects, that is applied without deliberate inference and therefore without the metacognitive awareness that would be required to inhibit it. The educational significance of this finding is substantial. It demonstrates that there exists a class of early childhood science misconceptions that is genuinely independent of sociocultural background, and for which the same instructional strategy, a single, vivid, dramatic disconfirmation, is equally effective across groups. The size-strength experiment constituted the most cognitively memorable moment of the entire intervention for children in both groups, producing spontaneous expressions of surprise and immediate revision. This outcome is consistent with Smith et al. (1993) whose argument that misconceptions should not always be characterised as flawed beliefs requiring replacement but as knowledge elements that are productive in many contexts and that require contextual refinement, specifically, the explicit marking of the boundaries of the contexts in which the p-prim applies versus those in which it does not, rather than elimination.

The most theoretically consequential finding of this study is not the accuracy of post-test responses but the systematic divergence between the two groups in the qualitative depth of the conceptual change those responses reflect. Across all four misconception targets, Group A children who achieved correct post-test responses did so predominantly through empirical reasoning, anchoring their revised beliefs in specific episodic observations (“I saw it stick”, “The small one held more”) rather than in reconstructed causal schemas. Group B children, whose pre-existing partial frameworks provided the scaffolding necessary to integrate new experimental evidence into a generalisable account, more frequently produced protoscientific responses that invoked causal mechanisms (“The force passes through the plastic”, “Size doesn’t make it stronger, it’s the power inside”).

This depth dimension is precisely what Vosniadou's (2013) framework theory predicts. When new information is encountered by a learner without an adequate framework to receive it, it tends to be assimilated as an isolated empirical fact rather than restructured into a

conceptually coherent account. The result is what Vosniadou (2013) terms a synthetic model. A conceptual structure that integrates some features of the new scientific information with the remnants of the prior framework, producing an internally inconsistent but locally functional epistemic state. The Group A post-test response to the permeability target, “Yes, I saw it pull through the water”, is a paradigm case of a synthetic model. The correct behaviour is registered, the animistic causal attribution has been abandoned, but the underlying physical mechanism (magnetic field permeability) has not been reconstructed. The response functions correctly in the immediate post-test context but lacks the conceptual architecture that would make it generalisable to novel instances of the same principle, such as magnetic field action through glass or through air gaps of varying widths.

This interpretation carries a significant pedagogical implication that the accuracy data alone would obscure. A one-week IBL intervention is powerful enough to eliminate animistic causal language and produce universally correct post-test responses on three of four misconception targets in Roma preschool children who had arrived at school with no prior domain-specific knowledge. This is a remarkable pedagogical achievement. It is not, however, equivalent to the achievement of proto-scientific conceptual understanding, and treating it as such would risk overstating the durability of the gains. The empirical-reasoning state that Group A children occupy at post-test is best understood as an intermediate stage in the conceptual change process, more advanced than the animistic attribution with which they began, less advanced than the causal-explanatory understanding they have not yet reached, rather than as a final destination. The instructional implication, as diSessa (1993) and Smith et al. (1993) both suggest, is that this intermediate state is not a failure but a productive resource. It provides the empirical foundation on which sustained, scaffolded instruction can build the causal schema that a single week of intervention cannot fully construct.

Among the four misconception targets, the all-metals misconception proved the most resistant to change, showing incomplete revision in Group A even among consistent attenders and complete non-revision among those whose attendance was disrupted. This outcome is consistent with Vosniadou’s (2013) prediction that framework-embedded beliefs, beliefs that are not isolated errors but expressions of a broader inferential principle, in this case the principle that perceptual surface features reliably predict physical behaviour, are more resistant to revision than local factual errors or universal p-prims, because revising them requires not just counter-evidence but the construction of an alternative framework principle that can displace the existing one. A single empirical disconfirmation, showing that the aluminium strip does not attract, provides counter-evidence against the specific prediction but does not, by itself, supply the alternative framework principle (ferrous/non-ferrous composition as the relevant criterion) that would generalise the revision to future instances. Multiple exposures, distributed across several lessons and including explicit scaffolding of the compositional distinction, are required for this generalisation to occur. The attendance effect observed within Group A directly confirms this prediction that children who attended all six lessons, and who therefore received the full distributed scaffolding of the material classification sequence, achieved the revision; children who missed even one lesson in the classification sequence did not.

This finding has a direct pedagogical implication for the teaching of magnetism in settings with irregular student attendance. The material classification target is not merely the hardest misconception to revise; it is the most sequentially dependent, requiring cumulative exposure rather than a single decisive experience. In settings where irregular attendance is structurally likely, as it is in the Roma school communities documented in this study, the TLS design must anticipate this dependency and provide either redundant exposures across multiple lessons or catch-up mechanisms for children who miss sessions. The current TLS architecture concentrated the material classification work in Lessons 1 and 2; redistributing it across three or four lessons, or providing brief consolidation activities at the opening of each lesson, would reduce the vulnerability of the classification gain to single-session absence.

The permeability misconception stands apart from the other three targets in two respects that have not been adequately noted in the existing literature. First, as the pre-test data demonstrate, it was held by a minority of children in Group A rather than a majority, in apparent contrast to the general pattern of Group A underperformance, because the p-prim underlying the correct conception (“forces act at a distance and penetrate obstacles”) is at least as cognitively available as the p-prim underlying the misconception (“physical barriers prevent physical action”) and is independent of prior informal exposure to magnets specifically. Second, the misconception was held not only by a minority of children but by the majority of adult educators present during the intervention, confirming that the permeability

principle is counterintuitive not merely for preschool-age children but for the general adult population. This adult misconception finding is consistent with Bar et al. (1997). Their documentation of the cross-age stability of action-at-a-distance errors, and it carries a direct professional development implication refer to the permeability of magnetic fields through non-magnetic media is not a piece of specialist content knowledge that preschool teachers can be assumed to hold, and explicit inclusion of this concept in science content training for early childhood educators is necessary.

The overall pattern of findings generates a set of pedagogical implications that go beyond the simple conclusion that IBL works. First, IBL instruction that produces empirical reasoning, correct responses anchored in specific observations, should be understood as a necessary but not sufficient condition for durable conceptual change. Sustained instructional scaffolding that explicitly builds the causal language and the generalisable principles on which empirical observations rest is required to move children from the empirical to the protoscientific reasoning level, and this scaffolding takes more time than a one-week intervention can provide. Second, the differential resistance of different misconceptions to change, with the all-metals misconception requiring distributed, sequential exposure and the size-strength equation yielding to a single dramatic disconfirmation, argues for a deliberate sequencing of misconception-targeted activities based on their theoretical resistance profile rather than their intuitive accessibility. Third, the complete equivalence of the two groups on the size-strength target at post-test, a universal p-prim, equally held and equally revised, suggests that the pedagogical design principle of vivid, personally conducted empirical disconfirmation is universally effective regardless of prior knowledge level or sociocultural background, and should be prioritised in any IBL sequence that aims to serve heterogeneous classrooms equitably. Finally, the depth divergence between groups at post-test is not a deficit of the Roma children but a structural consequence of the differential prior frameworks they brought to the intervention. Addressing this divergence requires not remediation but extended, linguistically rich science engagement of the kind that informal home environments already provide for children from higher sociocultural backgrounds, and that schools serving Roma communities must therefore provide explicitly and systematically.

6. Conclusions

The present study set out to investigate three interrelated questions: whether Roma and general-population preschoolers differ in the prevalence and patterning of magnetism misconceptions at school entry; which specific misconceptions prove most resistant to conceptual change; and whether an IBL TLS produces qualitatively different conceptual change trajectories in the two groups. The findings answer all three questions in ways that carry both theoretical and practical consequence.

With respect to the first question, the study confirms that the two groups enter the preschool science classroom with not merely different quantities of prior knowledge but qualitatively different epistemic configurations. Roma children in Group A presented predominantly with an absent framework, no prior domain-specific conceptual structure for magnetism, and no prior contact with the relevant vocabulary, leading to animistic and magical causal attribution rather than the misconceptions in the strict theoretical sense that characterise children with partial prior knowledge. General-population children in Group B presented with a largely accurate partial framework, within which locally overextended heuristics, such as the extension of metallic appearance to non-ferrous metals, constituted the primary misconception configuration. These are pedagogically distinct starting points that a single aggregate accuracy score cannot distinguish, and their differentiation is the foundational contribution of this study's analytical approach.

With respect to the second question, the study provides a theoretically grounded ranking of misconception resistance. The all-metals misconception proved the most resistant to change, because it is embedded in a broader perceptual-associative framework principle, surface appearance as a reliable predictor of physical behaviour, that requires sustained, distributed instructional exposure rather than a single disconfirmation to displace. The animistic attribution and permeability misconceptions, by contrast, yielded to appropriately designed single-session interventions, because they are either categorically isolated errors or the product of competing p-prims of roughly equal cognitive availability. The size-strength equation occupied a unique position – universally held and universally revised, it responded dramatically to a single vivid empirical disconfirmation regardless of group membership, confirming its status as a phenomenological primitive whose contextual boundaries can be

marked through a single decisive experimental contrast.

With respect to the third question, the study's most theoretically significant finding is the consistent divergence between the two groups in the qualitative depth of the conceptual change achieved, even where surface accuracy was identical. Group A children predominantly achieved empirical revision, correct post-test responses anchored in episodic memory of observed experimental outcomes, while Group B children more frequently produced protoscientific responses that integrated new evidence into a generalisable causal account. This divergence is not a product of differential cognitive capacity but of differential prior framework availability, particularly children who arrive with a partial conceptual scaffold can integrate new evidence into it and produce explanatory generalisations; children who arrive without such a scaffold can acquire empirically correct responses but lack the conceptual architecture to embed them in a durable causal account. The intermediate empirical-reasoning state achieved by Group A children at post-test is therefore not a failure of the intervention but a genuine and theoretically predictable stage in the conceptual change process, one that represents a substantial advance over the animistic attribution with which those children began, and that constitutes the productive foundation on which further sustained instruction can build.

Three theoretical conclusions follow from these findings. First, the conceptual change literature's central distinction, between the mere replacement of an incorrect belief and the genuine restructuring of the framework within which it is embedded, is empirically observable at the preschool level, and is detectable through qualitative analysis of children's reasoning and justification in the diagnostic interview even among four-to-six-year-olds. The methodology developed in this study, the four-category reasoning type scheme applied to verbatim interview responses, provides a replicable instrument for making this distinction operational in future research. Second, the differential resistance of different misconceptions to conceptual change is theoretically predictable from the structure of the prior conception, and instructional design that takes this predictability seriously, allocating more distributed exposure to framework-embedded misconceptions and more vivid single-contrast experiences to universal p-prims, will be more effective than design that treats all misconceptions as equivalent targets. Third, the universality of the size-strength p-prim and its universal revision through identical instruction confirms that there exists a class of early science misconceptions for which sociocultural background is genuinely irrelevant to both the prevalence of the error and the effectiveness of the remedy, and identifying this class has important implications for equitable instructional design.

Three practical conclusions follow with equal weight. For teachers designing science instruction in socially heterogeneous preschool classrooms, the finding that animistic and magical causal attribution in children from marginalised backgrounds is not a misconception but a categorical gap requires a fundamental reorientation of pedagogical expectation. Such children do not need to have an incorrect belief corrected; they need to be introduced to the domain itself. IBL, with its foundation in direct physical experience rather than prior linguistic or declarative knowledge, is an appropriate vehicle for this introduction precisely because it does not presuppose the prior framework that school entry has not yet provided. For curriculum designers, the sequential dependence of the all-metals revision on distributed, multi-lesson exposure, and its corresponding vulnerability to single-session attendance disruption, argues for a TLS architecture that distributes material classification activities across multiple lessons and provides brief consolidation at the opening of each subsequent session. For teacher educators, the finding that the magnetic field permeability principle was held as a misconception by the majority of adult educators present during the intervention is not an incidental observation but a systemic signal. It means that early childhood science teachers cannot be assumed to hold correct conceptions of the phenomena they are asked to teach, and subject matter preparation in early science content, not merely in IBL pedagogy, must be a core component of both initial teacher education and continuing professional development.

The deepest implication of this study's findings is perhaps the simplest to state. A Roma child who arrives at kindergarten with no prior knowledge of magnets, no relevant vocabulary, and no conceptual framework for the physical phenomena under study, and who, if she attends consistently, departs a week later able to predict, classify, and provide empirical justifications for magnetic behaviour across three of four conceptual targets, has not merely learned some science. She has demonstrated that the cognitive machinery for conceptual change in physical science is available to her, fully and without deficit, and that what has been absent is not the capacity for scientific thought but the opportunity to exercise it. The task

for research is to continue mapping the conditions under which that opportunity can be reliably provided. The task for practice is to provide it.

Author Contributions: Both authors contributed equally to this work. D.K. and K.T.K. jointly conceived and designed the study. Both authors contributed to the development of the experimental setup, the instructional design, and the research methodology. Data collection was carried out collaboratively, and both authors contributed equally to data analysis and interpretation. The manuscript was written jointly by both authors, and both approved the final version for submission.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with internationally accepted ethical standards for research involving human participants. Ethical approval was obtained from the relevant school administration and educational authorities prior to data collection. Written informed consent was obtained from the parents or legal guardians of all participating students. Student participation was voluntary, and participants were informed of their right to withdraw from the study at any time without any consequences. All data were collected and analyzed anonymously, and strict confidentiality was maintained throughout the research process. No personal identifying information was recorded, and the study involved no physical, psychological, or social risk to the participants.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Bar, V., Zinn, B., & Rubin, E. (1997). Children's ideas about action at a distance. *International Journal of Science Education*, 19(10), 1137-1157. <https://doi.org/10.1080/0950069970191003>
- Barrow, L. H. (2000). Do Elementary Science Methods Textbooks Facilitate the Understanding of Magnet Concepts? *Journal of Science Education and Technology*, 9(3), 199-205.
- Betancur, L., Votruba-Drzal, E., & Schunn, C. (2018). Socioeconomic gaps in science achievement. *International Journal of STEM Education*, 5, 38. <https://doi.org/10.1186/s40594-018-0132-5>
- Bourdieu, P. (1977). *Outline of a Theory of Practice*. (R. Nice, Trans.). Cambridge: Cambridge University Press.
- Bourdieu, P. (1986). The Forms of Capital. In J. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241-258). New York: Greenwood.
- Cabra Hernández, H. W. (2025). Socioeconomic achievement gaps in science and social studies in Colombia: An assessment of factors affecting school achievement. *Frontiers in Education*, 10, 1658707. <https://doi.org/10.3389/feduc.2025.1658707>
- Cheng, M., & Brown, D. E. (2010). Conceptual Resources in Self-developed Explanatory Models: The importance of integrating conscious and intuitive knowledge. *International Journal of Science Education*, 32(17), 2367-2392. <https://doi.org/10.1080/09500690903575755>
- Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 61-82). Routledge.
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches*. Sage.
- DeCosta, P., Skinner, T. C., Sørensen, J. L., Krogh Topperzer, M., & Grabowski, D. (2023). Narrative and play-based interviewing – a framework for eliciting the perspectives of young children. *Qualitative Research in Psychology*, 20(3), 337-362. <https://doi.org/10.1080/14780887.2023.2195366>
- diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10(2-3), 105-225.
- Driver, R., & Easley, J. (1978). Pupils and Paradigms: a Review of Literature Related to Concept Development in Adolescent Science Students. *Studies in Science Education*, 5(1), 61-84. <https://doi.org/10.1080/03057267808559857>
- Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (2014). *Making Sense of Secondary Science: Research into children's ideas*. Routledge. <https://doi.org/10.4324/9781315747415>
- French, L. (2004). Science as the center of a coherent, integrated early childhood curriculum. *Early Childhood Research Quarterly*, 19(1), 138-149. <https://doi.org/10.1016/j.ecresq.2004.01.004>
- Ginsburg, H. P. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511527777>
- Graziano, M. S. A. (2022). A conceptual framework for consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, 119(18), e2116933119. <https://doi.org/10.1073/pnas.2116933119>
- Kähkönen, A.-L., Sederberg, D., Viiri, J., Lindell, A., & Bryan, L. (2020). Finnish secondary students' mental models of magnetism. *Nordic Studies in Science Education*, 16(1), 101-120. <https://doi.org/10.5617/nordina.5566>
- Kakana, D., Christidou, V., Kazela, K., & Valakosta, M. (2009). Teaching Magnetic Attraction to Preschool Children. *The International Journal of Learning: Annual Review*, 16(2), 115-128. <https://doi.org/10.18848/1447-9494/cgp/v16i02/46130>
- Kalogiannakis, M., Nirgianaki, G.-M., & Papadakis, S. (2018). Teaching magnetism to preschool children: The effectiveness of picture story reading. *Early Childhood Education Journal*, 46(5), 535-546. <https://doi.org/10.1007/s10643-017-0884-4>
- Kotsis, K. T. (2023). Alternative ideas about concepts of physics, a timelessly valuable tool for physics education. *Eurasian Journal of Science and Environmental Education*, 3(2), 83-97. <https://doi.org/10.30935/ejsee/13776>
- Kotsis, K. T. (2024). The Significance of Experiments in Inquiry-based Science Teaching. *European Journal of Education and Pedagogy*, 5(2), 86-92. <https://doi.org/10.24018/ejedu.2024.5.2.815>

- Liljedahl, P. (2011). The theory of conceptual change as a theory for changing conceptions. *NOMAD Nordic Studies in Mathematics Education*, 16(1-2), 5-28. <https://doi.org/10.7146/nomad.v1i1-2.148290>
- McCormick, S. R., Brennan, P. O., & Yassa, J. G. (2000). Magnets and children – an attractive combination?. *BMJ (Clinical research ed.)*, 321(7259), 514. <https://doi.org/10.1136/bmj.321.7259.514>
- Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Science Achievement Gaps Begin Very Early, Persist, and Are Largely Explained by Modifiable Factors. *Educational Researcher*, 45(1), 18-35. <https://doi.org/10.3102/0013189X16633182>
- Naushabekov, Z., Maxutov, S., Mansurova, A., & Japashov, N. (2025). Effect of project-based learning on middle school students' conceptual understanding of mechanics: A quasi-experimental study. *European Journal of Science and Mathematics Education*, 13(3), 225-237. <https://doi.org/10.30935/scimath/16826>
- Ngazizah, N., & Fatmaryanti, S. D. (2025). Diagnosing and Remediating Misconceptions in Magnetism: An Integrated Conceptual Change and Simulation-Based Approach for Elementary Students. *Jurnal Penelitian Pendidikan IPA*, 11(12), 142-148. <https://doi.org/10.29303/jppipa.v11i12.13016>
- Piaget, J. (1952). *The origins of intelligence in children*. (M. Cook, Trans.). W. W. Norton & Company. <https://doi.org/10.1037/11494-000>
- Piaget, J. (1989). *The Child's Conception of the World: A 20th-Century Classic of Child Psychology*. Bloomsbury Academic.
- Ponizovsky-Bergelson, Y., Dayan, Y., Wahle, N., & Roer-Strier, D. (2019). A Qualitative Interview With Young Children: What Encourages or Inhibits Young Children's Participation? *International Journal of Qualitative Methods*, 18. <https://doi.org/10.1177/1609406919840516>
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211-227. <https://doi.org/10.1002/sce.3730660207>
- Ravanis, K. (1994). The discovery of elementary magnetic properties in preschool age: Qualitative and quantitative research within a piagetian framework. *European Early Childhood Education Research Journal*, 2(2), 79-91. <https://doi.org/10.1080/13502939485207621>
- Ravanis, K. (2022). Research Trends and Development Perspectives in Early Childhood Science Education: An Overview. *Education Sciences*, 12(7). <https://doi.org/10.3390/educsci12070456>
- Samara, V., & Kotsis, K. T. (2023). The use of new technologies and robotics (STEM) in the teaching of sciences in primary education: The concept of magnetism: A bibliographic review. *European Journal of Education Studies*, 10(2). <http://dx.doi.org/10.46827/ejes.v10i2.4652>
- Samara, V., & Kotsis, K. T. (2024b). Use of the artificial intelligence in teaching the concept of magnetism in preschool education. *Journal of Digital Educational Technology*, 4(2), ep2419. <https://doi.org/10.30935/jdet/14864>
- Samara, V., & T. Kotsis, K. (2024a). Research on kindergarten and primary school students' magnetic misconceptions and how to plan educational activities. *EIKI Journal of Effective Teaching Methods*, 2(2). <https://doi.org/10.59652/jetm.v2i2.124>
- Smith III, J. P., diSessa, A. A., & Roschelle, J. (1994). Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition. *Journal of the Learning Sciences*, 3(2), 115-163. https://doi.org/10.1207/s15327809jls0302_1
- Smolleck, L. (2011). Playing With Science: An Investigation of Young Children's Science Conceptions and Misconceptions. *Current Issues in Education*, 14(1). <https://cie.asu.edu/ojs/index.php/cieatasu/article/view/324>
- Sowmya, A. S., & Rani, R. N. (2025). Emerging interest in science during early childhood period and evaluating its effectiveness. *Scientific Reports*, 15(1), 38683. <https://doi.org/10.1038/s41598-025-22513-1>
- Varela-Aldás, J., Collay, W., Del-Valle-Soto, C., & Palacios-Navarro, G. (2026). Does the use of dedicated mobile devices in magnetism classes improve student learning? *Frontiers in Education*, 11, 1796723. <https://doi.org/10.3389/feduc.2026.1796723>
- Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, 4(1), 45-69. [https://doi.org/10.1016/0959-4752\(94\)90018-3](https://doi.org/10.1016/0959-4752(94)90018-3)
- Vosniadou, S. (2013). Conceptual change in learning and instruction: The framework theory approach. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 11-30). Routledge.