This article describes an empirical study that investigated the impact of different notational variations used in UML class diagrams on human comprehension. Despite UML being widely adopted as a standard modeling language, there exists a lack of consistency in the syntactic notations used across various sources like textbooks, research papers, documentation, and CASE tools when depicting UML class diagrams. Seemingly equivalent notational representations are chosen based on the personal preferences of authors or publishers, rather than being guided by research on which notations best support comprehension by the human readers and users of these diagrams.

The motivation behind this research was to determine whether there are significant differences in how well humans can comprehend UML class diagrams based on the specific notational variants used for representing core constructs like inheritance hierarchies, associations between classes, and cardinality constraints. Five common notational variations were identified, each having two alternative syntactic representations labelled as (a) and (b).

An experiment was meticulously designed using materials like a textual specification of a system, tutorial on UML notation, and worked examples. 34 novice subjects (students) and 5 expert subjects (industry professionals) participated. The key task involved presenting the subjects with a series of UML class diagrams, some correct and some with intentionally introduced errors, and having them identify whether each diagram correctly represented the given textual specification or not. The diagrams used different combinations of the (a) and (b) notational variants across the five notational categories like inheritance direction, inheritance arcs, association representation, association naming, and cardinality depiction. Response times and accuracy of identifying correct/incorrect diagrams were measured.

The results revealed some interesting findings that went against the initial intuitions of the researchers based on perceived notation complexity. For certain notations, the presumably simpler (a) variations did lead to significantly better accuracy in correctly determining if a diagram matched the specification. However, surprisingly for other notations, the more visually complex (b) variations resulted in better speed and accuracy when it came to identifying diagrams that had errors intentionally introduced related to that notation.

The authors hypothesize that this counterintuitive finding may be because when subjects encountered unfamiliar or unintuitive notational representations, they became more cautious and diligent in carefully examining the diagrams to identify potential mistakes or deviations from the specification. In contrast, with familiar simple notations, they may have felt overly confident in a cursory validation.

The expert subjects were also asked for their preferences and rationale regarding the notations. Their responses generally aligned with choosing the simpler (a) variations based on intuition of clarity, avoiding ambiguity, and following natural conventions like positioning generalized classes above specialized ones in inheritance hierarchies.

The study concludes that when selecting notations to depict UML diagrams in teaching materials, tools or actual practice, the choices should be guided by empirical evidence on which notational variants best support comprehension for the intended usage scenarios and tasks. While simpler, more intuitive notations may aid overall understanding, more complex notations may offer benefits for certain tasks like rigorous analysis to catch errors and deviations.

The authors recommend further empirical investigations in real-world contexts and over longer study periods to understand if the notations considered most suitable vary depending on whether the focus is learning UML, using it individually for documentation or analysis, or applying it in a collaborative team-based software development process.

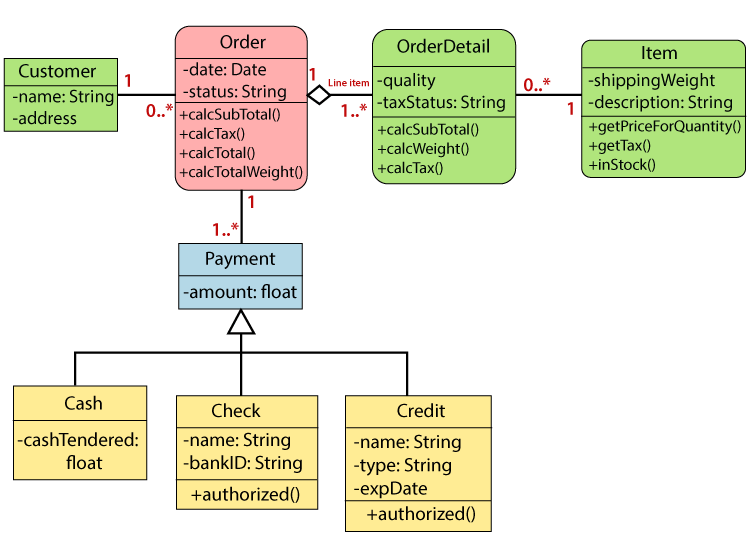
Overall, this work constitutes one of the earliest empirical studies to explicitly investigate the impact of seemingly superficial choices in syntactic notation design on the comprehensibility and effectiveness of UML diagrams as a communication and modeling tool for humans. It highlights the importance of making such choices based on evidence rather than individual preferences alone.

FIG 1 : SALES ORDER SYSTEM