Computers in Industry 165 (2025) 104232

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.sciencedirect.com/journal/computers-in-industry

Streamlining Assembly Instruction Design (S-AID): A comprehensive systematic framework

Mirco Bartolomei [a], Federico Barravecchia [a], Luca Mastrogiacomo [a], Davide Maria Gatta [b], Fiorenzo Franceschini [a,*]

[a] Politecnico di Torino, Department of Management and Production Engineering, Corso Duca degli Abruzzi 24, Torino 10129, Italy
[b] Leonardo S.p.A, Viale dell'Aeronautica, Pomigliano d'Arco 80038, Italy

ARTICLE INFO

Keywords:
Assembly Instruction Design
Information Processing
Worker Assistance
Assembly
Work Instruction

ABSTRACT

Assembly instructions are detailed directives used to guide the assembly of products across various manufacturing sectors. As production processes evolve to become more flexible, the significance of assembly instructions in meeting rigorous efficiency and quality standards becomes increasingly pronounced. Nevertheless, the development of assembly instructions often remains unstructured and predominantly dependent on the experience or personal skills of the designer. This paper aims to address these issues by pursuing three main goals: (i) deciphering the assembly process and the information that characterizes it, thereby providing a taxonomy of instruction constituents; (ii) presenting a framework to assess the various formats in which such information can be communicated; and (iii) introducing a step-by-step method, named *S-AID*, which offers a consistent methodology for designers during the instruction design phase. Overall, this research provides a rigorous taxonomy of the building blocks of assembly instructions and defines their relationships with various instruction formats. Furthermore, by proposing a systematic design method, this works aims to address the redundancy and inconsistency commonly encountered in traditional instruction design processes. The proposed methodology is illustrated using a real-world case study involving the assembly of a mechanical equipment. Finally, the effectiveness of the *S-AID* method was evaluated quantitatively through comparative analysis with other instruction sets, focusing on metrics such as process failures, assembly completion time, and perceived cognitive load.

1. Introduction

The manufacturing industry is currently undergoing a significant transformation. Traditionally dominated by mass production methodologies, the sector is now increasingly embracing a paradigm shift towards mass customisation (Barravecchia et al., 2023; Villani et al., 2018). This shift demands more flexible production processes to manage diverse product varieties (Capponi et al., 2024; Gervasi et al., 2020). While crucial for maintaining competitiveness, these adaptable processes can often be onerous and complex to manage for the workforce. By guiding operators through unfamiliar and complex assembly processes, well-designed assembly instructions have the potential to mitigate the likelihood of errors and boost efficiency (Fiorentino et al., 2014). This is particularly important in sectors where high competitiveness or strict quality standards prevail (Franceschini et al., 2019).

While the need for comprehensive assembly instructions is widely recognised, much of the existing research in this area focused on evaluating different formats for instruction delivery (Dorloh et al., 2023). Numerous studies examined the impact of different instruction presentation formats on operator performance, exploring ways to optimise the effects of instructions (Arguel and Jamet, 2009; Van Genuchten et al., 2012). Further research investigated tailoring instructions to match the operator's experience level or the task's complexity (Radkowski et al., 2015). Additionally, comparisons among various instruction delivery tools—including augmented reality (AR), paper-based, and video methods—showed significant differences in performance outcomes. Notably, AR instructions showed to significantly reduce error rates when compared to video-based instructions (Loch et al., 2016), with these beneficial effects increasing with task complexity (Alessa et al., 2023).

Despite existing studies, there remains a lack of a systematic, adaptable, and reproducible methodology for designing assembly instructions, highlighting a significant gap in the literature (Geng et al.,

* Corresponding author.
E-mail address: fiorenzo.franceschini@polito.it (F. Franceschini).

https://doi.org/10.1016/j.compind.2024.104232
Received 23 January 2024; Received in revised form 9 July 2024; Accepted 15 December 2024
Available online 24 December 2024

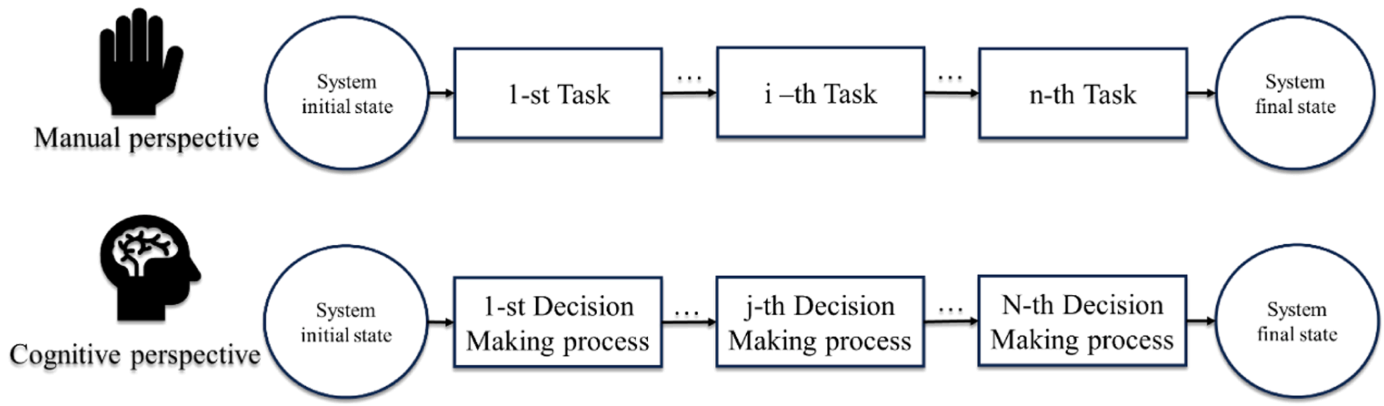


Fig. 1. Representation of assembly sequence from a manual and cognitive perspective.

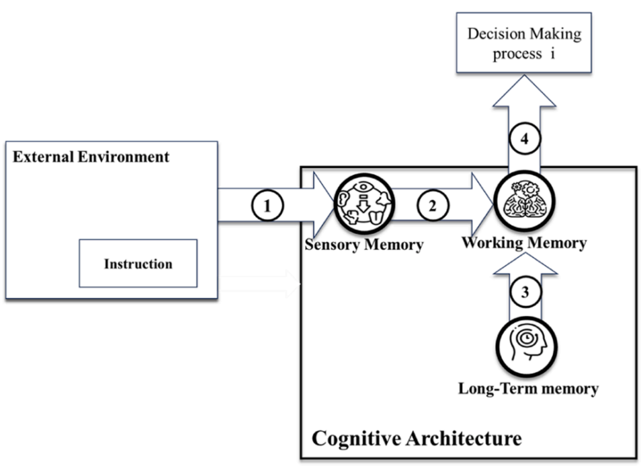


Fig. 2. Exemplification of the human Information Processing (HIP) model (Atkinson and Shiffrin, 1968; Ganier, 2004; Wickens et al. 2021). Symbol legend: 1 = Sensory memory raw stimulus processing; 2 = Working memory information element extraction; 3 = Working memory schemas retrieval; 4 = Working memory action selection and planning.

2020; Laviola et al., 2024). This study therefore seeks to address this gap by introducing a *Streamlined Assembly Instruction Design (S-AID)* method, aimed at enhancing the adaptability and efficiency of assembly instructions. This method provides a systematic approach to defining the number of instructions required to accurately represent an assembly process, determining the content of each instruction, and selecting the most appropriate formats for their presentation.

The remainder of this paper is structured as follows. Section 2 analyses the assembly process from a cognitive perspective and provides a taxonomy of the constituent elements of instructions. Section 3 gives an overview of instruction formats and provides a framework for representing their relationship to the instruction building blocks defined in Section 2. Section 4 provides a map for assessing instruction formats. Section 5 presents the proposed *S-AID method*. Section 6 presents an empirical investigation of the S-AID method. Finally, the discussion and conclusion section discusses the contributions, limitations and possible future research directions.

2. Background

An assembly sequence is a detailed process that outlines the steps required to assemble various parts into a complete product. It specifies the exact order in which parts should be combined in an assembly process and the correct way to integrate them (Hu et al., 2011). The steps in the assembly sequence, referred to as tasks, transform a system from its initial state to its final state, resulting in a subassembly or finished product. However, such a description provides only a partial understanding of the assembly sequence. Before any coupling occurs, components must be recognised, grasped, oriented, and correctly positioned in their mounting location (Boothroyd and Alting, 1992). Each of these manual actions results from a cognitive decision-making process. In particular, the operator uses the information at his disposal to determine and programme the actions necessary to complete the assembly sequence (Ganier, 2004).

Viewing the assembly sequence from this perspective reveals it as a chain of cognitive decisions, each determining the course of manual actions (see also Fig. 1). In scenarios where the operator is familiar with the process, these decisions are made automatically. Conversely, for unfamiliar processes, the operator requires supporting information to guide decision-making.

2.1. Human information processing model

Instructions are an indispensable guide for inexperienced operators or those faced with unfamiliar tasks. Thus, by comprehending the ways in which human beings process information for decision-making, the necessary content for instructions can be determined.

The segment of the human mind responsible for decision-making is known as the cognitive architecture and is subject of interest in cognitive science (Del Missier et al., 2013; Patterson et al., 2014). This architecture comprises three types of memory: *sensory memory, working memory, and long-term memory* (Sweller et al., 2019). The process by which external information is processed by the cognitive architecture and used to conduct decision-making processes can be outlined as follows (see also Fig. 2):

1) *Sensory memory raw stimulus processing:* sensory memory acts as an initial gateway for external stimuli, capturing information through sensory organs. It processes a sequence of sensory inputs by converting them into electrical impulses. These impulses are then further transformed into raw information that can be processed by working memory (Atkinson and Shiffrin, 1968). Assembly instructions are part of the external stimuli that an operator receives while executing the assembly.
2) *Working memory information extraction:* working memory processes incoming sensory information, comprising colours, shapes, sounds, and haptic sensations (Wolfe, 1994). Working memory scans this continuum to extract discrete information chunks, known as information elements, which guide decision-making processes to achieve predetermined task goals (Miller, 1956). To illustrate, when determining the optimal positioning of a component within a subassembly with multiple potential orientations and mounting positions, two distinct information elements are required: one defines the correct

Table 1
Taxonomy of AFs proposed by Van Holland and Willem (2000). Symbol legend: D = Dynamic information. S = Static information.

Category	*Assembly Feature*	*Type*	Description
Handling	*Feeding*	*S*	*Feeding* is the way the component is supplied into the assembly system, e.g. by trays or feeders, and the predefined position and orientation of the component.
	Fixturing	*D*	*Fixturing* represents the way the component must be fixed in jigs or fixtures.
	Gripper	*D*	*Gripper* represents how the product is gripped from the supply position, i.e. the way it is fitted into the tool in use.
	Grasping areas	*S*	*Grasping Areas* represent the allowed contact surfaces between the tool used to move the part and the part itself.
Connecting	*Involved form feature type*	*S*	*Involved form feature type* represents the shape of the faces that come into contact during fastening.
	Final position	*S*	*Final Position* represents the position and orientation assumed by the subassembly at the end of the assembly operation.
	Insertion position	*S*	*Insertion Position* represents the position and orientation assumed by the sub-assembly in the initial stages of insertion into another component.
	Insertion path	*D*	*Insertion Path* represents the trajectory linking the final position and insertion position of the assembly.
	Tolerances	*S*	*Tolerances* represent the dimensional tolerances for connecting a component to a sub-assembly.
	Contact areas	*S*	*Contact Areas* represents to the surfaces of the component and sub-assembly that come into contact during the insertion path;
	Internal freedom of motion	*D*	*Internal freedom of motion* concerns the range of movements allowed to the component within the sub-assembly once it is fastened.
	Geometric refinements	*S*	*Geometric Refinements* are special refinements to facilitate assembly, e.g. rounding and chamfering.

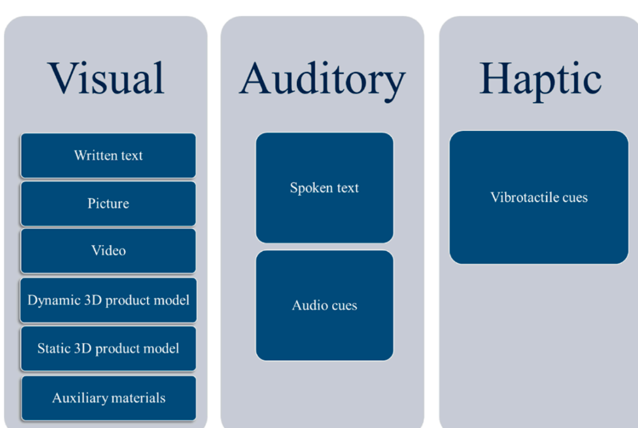


Fig. 3. Taxonomy of single instruction formats.

component orientation, and the other specifies the appropriate mounting position. Working memory analyses sensory perceptions to identify these elements, thereby guiding decision-making to achieve task goals.

3) *Working memory schemas retrieval:* due to its limited capacity, working memory can only process a few information elements simultaneously, typically ranging from five to seven (Sweller, 1994). However, learning allows aggregation of elements into schemas stored in long-term memory (Chandler and Sweller, 1991). When needed, working memory retrieves these schemas, each loading as a single information element (Chandler and Sweller, 1991). From an assembly perspective, this implies that if the process is familiar, working memory can retrieve all necessary information from long-term memory schemas for decision-making.
4) *Working memory action selection and planning*: based on the external information and past experiences stored in the long-term memory, the operator decides and plans the manual actions required to perform the assembly (Ganier, 2004).

The information processing model offers invaluable insights for instruction design. To develop effective instructions, designers must analyse the decision-making processes within the assembly sequence to identify critical information elements. Subsequently, designers must differentiate between those elements that are likely to be familiar to the operator from previous experience or directly affordable from the component to be assembled (Norman, 2002), and those that are not. This allows the essential information required for the task to be defined. Finally, the instructions should be organised to contain a concise set of these elements.

2.2. Information elements taxonomy

The previous section relates the content of instructions to the notion of information element derived from cognitive science. As a result, for instruction design to be both systematic and scalable, it is necessary to search for the basic information in manufacturing that can play the role of information element.

To this end, the concept of *Assembly Features (AFs)* is herein introduced. *AFs* encompass the geometrical and non-geometrical properties of product components that can influence the modelling, planning, and execution of the assembly process. These features have significant mechanical and functional implications, as they are the elements that enable the final product to perform the functions for which it is designed (De Fazio et al., 1993). Notwithstanding the mechanical nature of AFs, the operator's cognitive architecture utilizes these attributes as essential information elements for understanding the assembly sequence (Wang et al., 2021). This establishes a vital connection between the manufacturing domains and the informational domain.

The taxonomy proposed by Van Holland and Willem (2000) (see Table 1) is considered in the *S-AID method*, as it combines flexibility and ease of use. The chosen taxonomy encompasses a wide range of AFs categories, allowing for flexibility in representing the diverse assembly processes found in manufacturing settings. Additionally, it features a concise set of items, prioritizing simplicity and user-friendliness.

Van Holland and Willem (2000) distinguish between two categories of *AFs*:

- *Handling Features*: refer to the collection of information necessary for manipulating the component involved in the assembly process. Handling AFs range from the initial position of the component when it is fed into the work area to the way the component is transported to set up assembly.
- *Connecting Features:* denote the compilation of information required for affixing the component in the appropriate assembly location. Connecting AFs span from the surfaces involved in fastening to the specification of the component's orientation prior to and following assembly.

AFs can represent either *static information*, i.e. the spatial relationships between the elements involved in the assembly and their tolerances, or *dynamic information*, which can only be fully rendered by showing the evolution of the system over time. While the taxonomy of

Joining Flange

Instructions for Securing Flange :

Positioning of Oval Flange : Place Oval Flange on the upper surface of the Square Flange such that it aligns with the four holes furthest from itself.

Orientation of Square Flange: Ensure that Oval Flange is positioned so that its flat face is facing upwards.

Bolts Commissioning: Retrieve two long bolts from the component feed tray.

Bolts Positioning: Insert the two bolts into their respective holes located on the surface of Oval Flange.

Nuts Commissioning: Retrieve two nuts from the component feed tray.

Tightening: Secure the nuts to the inserted bolts by rotating them clockwise until fully tightened.

Fig. 4. Written text instructions describing the assembly of a flange using two nuts and two bolts.

AFs allows for a comprehensive description of an assembly process, not all AFs have to be provided to the operator through instructions. Information requirements are herein redefined as the subset of AFs that the operator cannot derive from previous experience and therefore must be provided by instructions.

3. Instruction formats and assembly features

The variety of formats for presenting AFs within instructions is considerable, with each approach having its own potential and limitations. An overview of the different instruction formats and the AFs they can represent is given in the following sections.

3.1. Instruction formats

Instruction formats are characterised by two dimensions: (i) the sensory channels used to deliver information, and (ii) the presence of single or multiple types of media. A taxonomy of individual formats, distinguished by the engaged senses is provided in Fig. 3. Instruction formats span from visual to auditory, and haptic modalities. Senses such as taste, and smell are used by operators as a means of detecting hazards during production. However, in the manufacturing context, they are not currently used as conventional means of communicating instructions.

The diverse types of instruction formats are:

- *Written text:* this format can describe all AFs (Bieger and Glock, 1985). However, to correctly understand a textual instruction, its symbols undergo a semantic processing by the cognitive architecture (Baddeley, 1979). The semantic processing has two main effects: it increases cognitive load by overloading working memory, and when numerous AFs are described, it may hinder timely interpretation of the text. As an illustration, a joining operation, which entails the positioning of a component on a sub-assembly, the specification of its orientation, and the completion of the operation with a fastening, can be cumbersome to convey using textual instructions. An example of textual instruction describing a joining task is presented Fig. 4.
- *Picture:* provide a realistic representation of the system's state during assembly. This static visual information is explicit and does not require interpretation, allowing operators to easily extract the necessary details to perform the task at hand. As a result, pictures are generally timelier than text. However, due to their static nature, they can represent dynamic information only in an implicit way (Bieger and Glock, 1985, 1986). An example of a picture describing an assembled product is shown in Fig. 5A and Fig. 5B.
- *Video:* can dynamically depict the assembly process, providing a complete representation of the required actions (Hegarty, 2012). The sequential presentation of information in videos necessitates the viewer to complete the entire presentation to achieve full comprehension, which makes them less time-efficient than pictures. To illustrate, for simple tasks like identifying the component to be assembled, pictorial representations are more immediate than videos. Conversely, videos excel in guiding detailed motor tasks, such as fastening and grippers manipulation, where demonstrating precise movements and actions is crucial (Höffler and Leutner, 2007).
- *Static 3D product model:* with augmented reality devices 3D models of assembly components can be holographically displayed, overlaying onto reality. These static holographic representations share the representational capabilities of pictures but require less cognitive processing. Their three-dimensional nature allows operators to grasp object details, aiding mental visualization of the assembly system (Huk, 2006). An example of a static 3D product model is displayed in Fig. 5C.
- *Dynamic 3D product model:* a dynamic 3D product model can be created by animating a static 3D product model, which can then be presented in a video-like format. This enables the evolution of the assembly system to be represented. Such dynamic 3D models can be presented in an AR environment and share both the representational capabilities and limitations of videos.
- *Auxiliary Materials:* are a collection of visual aids such as arrows, circles and parts lighting designed to display a restricted set of AFs in an immediate and straightforward manner (Tversky et al., 2000). An example of auxiliary materials displaying the movement to secure a bolt is represented in Fig. 5D and Fig. 5E. It is important to note that these formats do not convey meaning per se; rather, they are dependent on external contextual information. For example, an arrow designated to indicate an object to be grasped is incapable of conveying information without an image representing the object

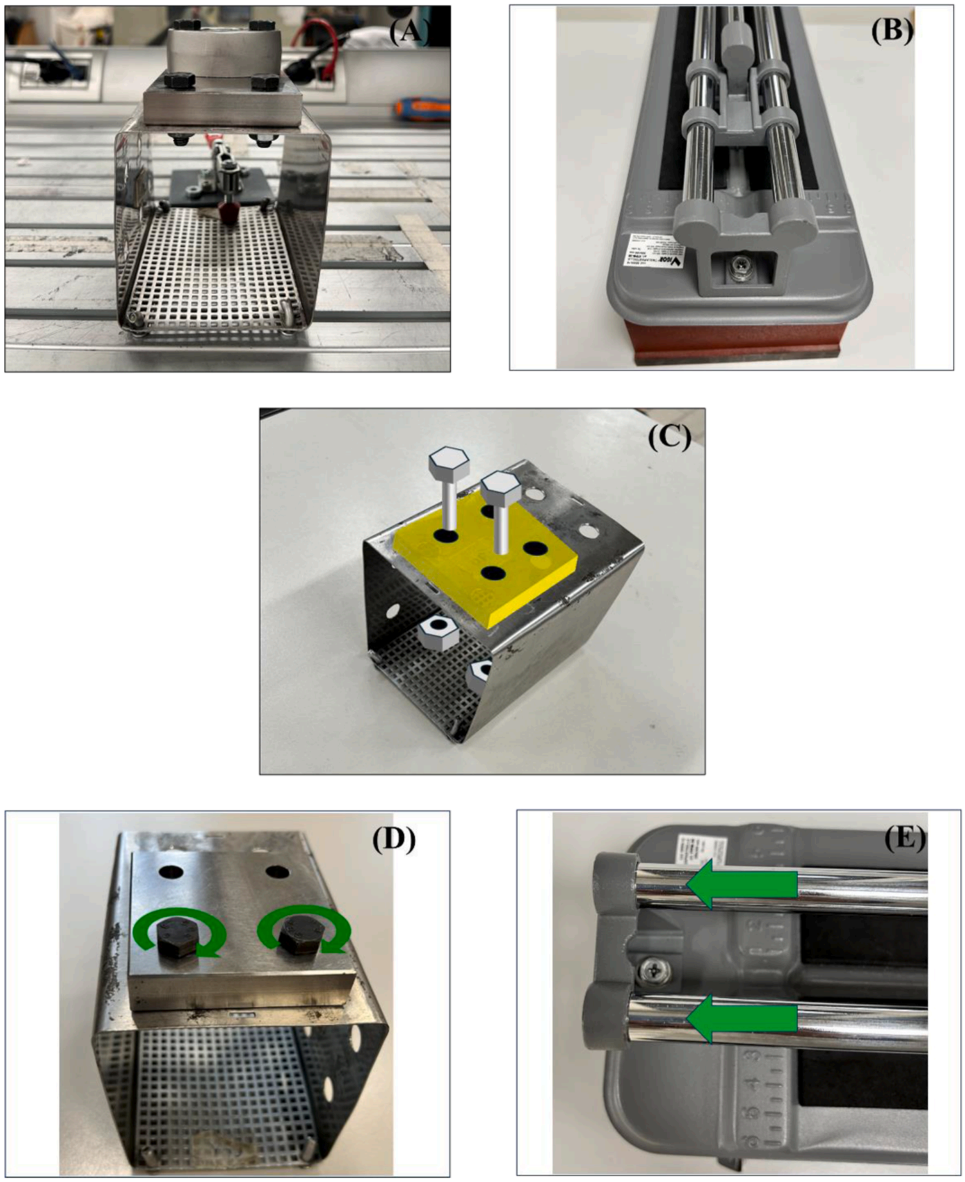


Fig. 5. Examples of instruction formats - (A) Picture showing the semi-assembled state of a mechanical component; (B) Picture showing the final assembled state of a tile cutter; (C) Static 3D product model describing the positioning of a component and the insertion of two bolts; (D) Auxiliary material showing the fastening of two bolts; (E) Auxiliary material showing the path for inserting two rods.

itself. For this reason, auxiliary materials are generally used within multimedia formats in combination with pictures or videos. This analysis considers them to be equivalent to independent formats for two main reasons. Firstly, the integration of AR allows the representation of auxiliary materials superimposed on the reality, thereby eliminating the need for additional context and enabling auxiliary materials to be presented independently. Secondly, an understanding of the relative strengths and weaknesses of these formats allows for their informed integration into multimedia formats.

- *Spoken Text:* the provision of textual information in audio mode can be utilized to describe the AFs required to perform an assembly. Spoken text and written text are of the same nature and suffer from the same weaknesses in terms of timeliness. In addition, for long and complex texts, the audio format may underperform the written text. Information is retained in the working memory for a short time unless it is continuously refreshed. While this rehearsal is straightforward for written text, it becomes more challenging for spoken text when the listener cannot control the pace of the information delivery (Dunham et al., 2020).
- *Audio Cues:* are unique sounds used to convey specific information or commands to a user. These can range from simple tones or beeps to more complex sound sequences. Their main use is in assisting humans in manipulating objects in AR and virtual reality (VR) environments (Zahariev and MacKenzie, 2007). However, the same approach may be extended to real manufacturing contexts. To illustrate, confirmation cues can be employed to provide immediate feedback on the completion of a task, such as part positioning or grasping (Canales and Jörg, 2020).
- *Vibrotactile Cues:* they direct the motor actions to be performed by the operator through continuous feedback provided by means of haptic stimuli (Sigrist et al., 2013). Vibrations are provided in the form of waves and can vary in parameters like amplitude, frequency, waveform, and duration (Han and Schulz, 2020). Varying these parameters enables continuous feedback to be generated regarding the

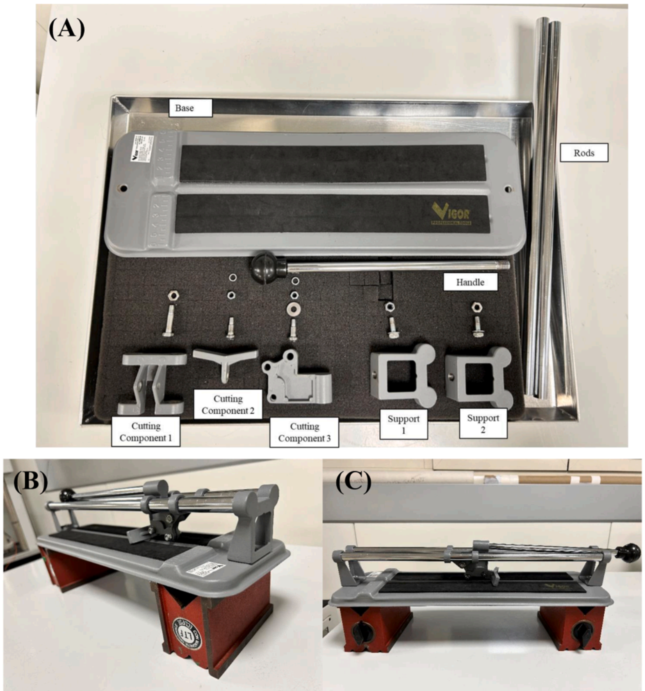


Fig. 6. Tile cutter used as an example for evaluating picture format representational ability - (A) Tile cutter components with their respective identifiers; (B) Assembled tile cutter from a lateral view; (C)Assembled tile cutter from a frontal view.

proximity to the location of a component to be gripped, or the correct position of the part to be assembled (Arbeláez et al., 2019).

- *Multimedia instructions:* multimedia instructions integrate multiple formats to capitalize on their respective strengths and provide comprehensive representational capabilities. Nevertheless, the effectiveness of multimedia instructions is contingent upon the complementary use of the constituent formats (Ayres and Sweller, 2014; Sweller and Chandler, 1994). Merely duplicating the same information across different formats can have negative effects on instruction processing (Kalyuga et al., 1999). For instance, the use of an arrow to indicate the position of a part to be grasped and the simultaneous provision of textual description in written form within the same instruction can result in a state of cognitive overload and a reduction in the clarity of the instruction. Common examples of multimedia formats used for presenting instructions include video and text (Lee and Shin, 2012), video and audio (Lee and Shin, 2012), picture and auxiliary materials (Heiser and Tversky, 2006), picture and text (Schnotz, 2014).

3.2. Representational ability framework

Based on the previous analyses, a comprehensive framework has been deployed to effectively associate AFs with different instruction formats. This framework enables designers to identify the specific information that each instruction format can represent, making it a valuable tool for identifying the formats capable of representing the information elements required by a task.

The construction of the framework is based on the following premises:

- The potential of each instruction format was evaluated by considering its individual and independent form. As an example, in traditional assembly instructions, pictures are often intended as multimedia format combining technical drawings or photographs with additional elements such as text, dimensioning, arrows, or markers indicating an area of interest. In contrast, in this study, a picture strictly refers to the depiction of the components involved in the task, alone or in their assembly context.
- An instruction format was deemed capable of depicting an AF only if it could provide a complete and explicit representation of the feature. As a practical example, an image capturing the final assembled state of a subassembly does not provide explicit guidance on the transition from its initial disassembled state. Consequently, the operator must infer this solely based on his skills or previous experience. For this reason, formats of a static nature have been considered uncapable of representing AFs which require the representation of the evolution of the system to be fully conveyed.

A case study illustrating the process of assembling a manual tile cutter serves as a representative example. Fig. 6 depicts the manual tile cutter, both in its assembled state and in its disassembled form. Table 2 provides a practical illustration for determining which AFs can be effectively represented by a picture. Table 3 presents the framework, which was developed by applying the same evaluation criteria used for the picture format to other instruction formats. Some examples of evaluations conducted for the other formats are discussed below:

- *Written and spoken texts:* are effective for describing both static and dynamic AFs due to the flexibility of verbal representation.
- *Audio and vibrotactile cues*: provide straightforward sensory feedback for basic spatial tasks, such as guiding the placement of components in assemblies or identifying the correct part to be assembled from the feeding tray. However, providing sensory feedback to guide complex manual actions such as fixturing or joining operations is not a viable option.
- *Videos and dynamic 3D models:* can convey both static and dynamic information. To illustrate, these formats can display how tools and parts interact, including the AFs related to gripping areas and component insertion paths. In addition, videos and dynamic 3D models can visually convey AFs of degree of freedom by showing the movement and allowable directions of components post-assembly. However, without textual integration, these formats may not adequately convey necessary assembly tolerances.
- *Auxiliary materials:* examples of auxiliary materials such arrows and circles enhance instruction by indicating component positions and highlighting specific regions on component surface, thereby representing AFs like grasping and contact areas. Furthermore, while these materials can indicate directions for simple insertions and degrees of freedom, they fall short in fully and explicitly representing complex AFs such as gripper and fixturing. Being non-verbal, they also cannot explicitly convey tolerances.

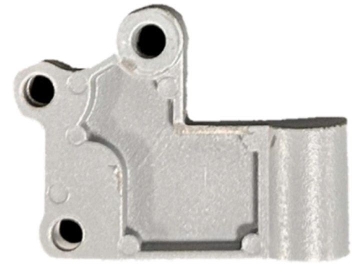
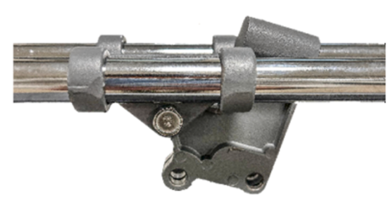
For the sake of brevity, multimedia formats (see Section 3.1) are not directly included in Table 3. Multimedia formats encompass the representational capabilities of the individual constituent formats. Therefore, the framework enables users to readily deduce the representational capabilities of a multimedia format by aggregating the AFs represented by its constituent formats.

4. Framework for instruction formats assessment

From the analysis of the instruction formats, it emerges that while some provide detailed information, their complexity can make them difficult to understand quickly, potentially affecting timely comprehension. Conversely, formats that prioritize immediacy tend to be less informative, lacking the details required for thorough understanding and execution.

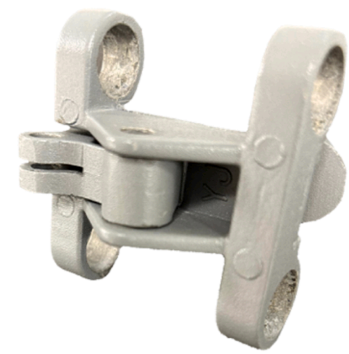
In the design of effective instructions, it becomes essential to consider these distinct format characteristics, ensuring they align appropriately with those of the task for which the instruction has been created.

Table 2
Evaluation of AFs suitable for pictorial representation. Symbol legend: N.A. = Not Applicable; the image cannot fully and explicitly represent the corresponding AF.

Assembly Features	Representable by means of a Picture format?	Motivation	Example of AF conveyed with a Picture format
Feeding	*Yes*	A picture of the component helps quickly identify and distinguish it from others in the feeding zone. The adjacent image serves as a simple reference for operators to recognize cutting component 3.	
Fixturing	*No*	Given its dynamic nature, fixturing cannot be fully captured through a static picture.	*N.A.*
Gripper	*No*	A picture is insufficient to capture the full dynamic process of fixing a component in a tool and then moving it to the assembly area.	*N.A.*
Grasping areas	*No*	Representing *Grasping areas* requires the identification of specific surface regions suitable for grasping. In a pictorial representation, supplementary graphic elements like colour coding are essential to distinguish these areas effectively.	*N.A.*
Involved form feature type	*Yes*	This AF involves spatial and static information since it requires recognizing the shapes of contacting faces during connection. For instance, the adjacent image highlights the shapes involved in assembling the support on the base, aiding in identifying the mounting position.	
Insertion position	*Yes*	*Insertion position* refers to the initial spatial orientation and position of a component before assembly. Its static nature makes it apt for pictorial capture. The adjacent image depicts the alignment of the support prior to be fixed on the base.	
Final position	*Yes*	The *Final position* pertains to the component's arrangement upon assembly completion. Its inherently static characteristics make it suitable for pictorial representation. The adjacent image represents the *final position* of the cutting component 3 by depicting the state of the subassembly after fastening completion.	
Insertion path	*No*	This AF requires the dynamic trajectory of a component during insertion to be represented. For this reason, it cannot be fully captured by a static picture.	*N.A.*
Tolerances	*No*	*Tolerance* details are often nuanced and require numerical or textual information, making them unsuitable for representation solely through a picture.	*N.A.*
Contact areas	*No*	This AF involves identifying regions on component surfaces that directly contact during assembly. Additional graphic elements within a component picture are crucial to distinguish these areas from others not in direct contact.	*N.A.*
Internal freedom of motion	*No*	The *Internal freedom of motion* involves the dynamic aspects of component movement within the assembly, making a static picture insufficient for its representation.	*N.A.*

(*continued on next page*)

Table 2 (*continued*)

Assembly Features	Representable by means of a Picture format?	Motivation	Example of AF conveyed with a Picture format
Geometric refinements	*Yes*	*Geometric refinements* are the geometric details of the components involved in the assembly that enables coupling. The image beside represents a case where a *geometric refinement*, i.e. a round, of cutting component 3 allows it to be easily inserted into its mounting position on cutting component 2.	

To drive the identification of most suitable assembly instruction format for a specific task, two evaluation dimensions are introduced: *Representational Ability* and *Processing Efficiency*.

Representational Ability of the i-th format for the j-th task ($RA_{i,j}$) is computed as follows:

$$RA_{i,j} = \frac{ReAF_{i,j}}{TR_j} \qquad (i = 1, \quad ..., n); (j = 1, \quad ..., N) \tag{1}$$

Where:

- *n:* total number of single and multimedia instruction formats assessed by the instruction designer.
- *N:* total number of tasks forming the assembly process under consideration.
- *Represented Assembly Features ($ReAF_{i,j}$):* denotes the number of AFs, identified as information requirements for the j-th task, that the i-th format can represent.
- *Total Requirements (TRj):* is the number of AFs identified as information requirements for the task being analysed. $RA_{i,j}$ assesses whether the instruction format possesses the capability to present the AFs required for the analysed task. An i-th format with an $RA_{i,j}$ value of less than 1 is not able to represent the j-th task.

Processing Efficiency ($PE_{i,j}$) of the i-th format for the j-th task is computed as follows:

$$PE_{i,j} = \frac{PT_{i,j}}{TT_j} \qquad (i = \quad 1, \quad ..., n); (j = 1, \quad ..., N) \tag{2}$$

Where:

- *n:* total number of single and multimedia instruction formats assessed by the instruction designer.
- *N:* total number of tasks forming the assembly process under consideration.
- *Processing Time ($PT_{i,j}$):* is the time required for an operator to process the i-th instruction format equipped with the j-th task's information requirements. This value is derived from an assessment conducted by the instruction designer. This assessment requires careful consideration of the characteristics inherent in the instruction format, as well as the type of AFs demanded by the task.
- *Task Time (TT_j):* is the estimated time to complete the j-th task.

The $PE_{i,j}$ relates the time needed to fully access and understand the content of an instruction to the duration of the task. As an illustrative calculation example, consider a j-th task where the execution time is 60 s, while the $PT_{i,j}$ of an i-th instruction format presenting the information requirements for the j-th task is 180 s. According to Eq. (2), the resulting value of $PE_{i,j}$ is 3. From a practical standpoint, this means that the time required to access and understand the instruction is three times longer than the duration of the task.

Considering these two variables, it is possible to place each instruc-

Table 3
Matrix relationship between AFs and single media formats.

			Single Media Format								
			Written text	Picture	Video	Static product model	Dynamic product model	Auxiliary material	Spoken text	Audio cues	Vibrotactile cues
Assembly Features	**Handling Features**	*Feeding*	•	•	•	•	•	•	•	•	•
		Fixturing	•		•		•		•		
		Gripper	•		•		•		•		
		Grasping areas	•		•		•	•	•		
	Connecting Features	*Involved form feature type*	•	•	•	•	•	•	•	•	•
		Insertion position	•	•	•	•	•		•		
		Final position	•	•	•	•	•		•		
		Insertion path	•		•		•	•	•		
		Tolerances	•						•		
		Contact areas	•		•		•	•	•		
		Internal freedom of motion	•		•		•	•	•		
		Geometric refinements	•	•	•	•	•		•		

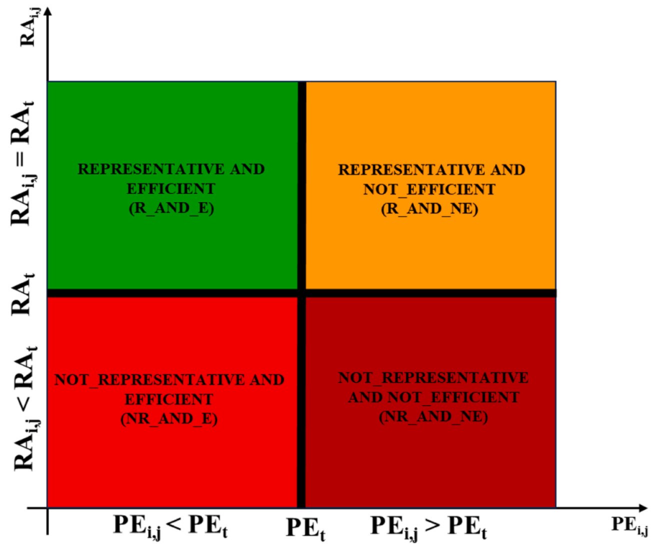


Fig. 7. Synthesis map of the instruction formats for each j-th task. Representational Ability ($RA_{i,j}$) VS. Processing Efficiency ($PE_{i,j}$).

tion format within the synthesis map outlined in Fig. 7. In the map, the x-axis represents the $PE_{i,j}$ variable, while the y-axis corresponds to the $RA_{i,j}$ variable.

The synthesis map displays four quadrants. The vertical threshold PE_t enables a distinction to be made between inefficient and efficient formats. The PE_t value is set by the instruction designer based on the emphasis on process efficiency. More rigorous time demands of the process necessitate a lower PE_t value. The horizontal threshold, on the other hand, represents the representational ability required to render the task at hand. The threshold value of RA_t is set equal to 1. Formats that fall below this threshold lack the necessary representational ability. Conversely, those that have an $RA_{i,j}$ value equal to 1 are considered capable of representing the task.

The content of the quadrants of the synthesis map can be described as follows:

- *The top-right quadrant*, named “Representative and Not Efficient”, contains instruction formats that have a robust representation capability but require an extended processing time. Such formats offer comprehensive and detailed depictions of the task’s information requirements, although they lack immediacy in information presentation.
- *The top-left quadrant*, named “Representative and Efficient”, contains instruction formats that provide a high degree of representativeness and require little processing time. These formats strike a balance between efficiency and effectiveness, providing clear and concise instructions.
- *The bottom-right quadrant*, named “Not Representative and Not Efficient”, contains formats that possess a low representational ability and require a long processing time. These formats may be inadequate in providing a precise and efficient depiction of the task’s requirements.
- *The bottom-left quadrant*, named “Not Representative and Efficient”, contains formats that exhibit a low representational ability and a short processing time. While these formats present information in a swift and straightforward manner, their simplicity restricts their ability to represent the task comprehensively.

5. S-AID step by step assembly instruction design

This section introduces a *Streamlined Assembly Instruction Design (S-AID)* method aimed at supporting the development timely and relevant instructions tailored to the operator’s experience level. The method guides designers through a step-by-step process, starting with the decomposition of the assembly sequence into subtasks, moving on to the identification of information requirements, and ending with an analysis of the capabilities and constraints of potential instruction formats. For every subtask, the final output of the method is a set of AFs that are required to convey the instruction, accompanied by an optimal format that can efficiently and effectively represent them.

The description of the method is accompanied by a case study that serves to illustrate the explanation of each step. The case study involves the assembly of a mechanical equipment (see Fig. 8A and Fig. 8B), whose parts are listed in Table 4.

5.1. Step 1: assembly process subdivision

In the first step, the assembly process is divided into different subtasks. This essential step enables the definition of the manual operations to be executed by the operator within each subtask. These operations are essentially what instructions are required to convey, making the subdivision of the process crucial for defining the content and boundaries of each assembly instruction.

The rationale used for subdivision is critical for two main reasons:

- *Alignment with AFs Representation Logic*: Task content must be consistent with the logic of AFs, which describe assembly at the component level. As a result, the assembly process is subdivided based on a component-oriented logic, with each main component assigned a task encompassing all required operations. Connectors, when present, are managed within the component’s assembly rather than being treated as separate tasks.
- *Balanced Operation Count*: Every segment resulting from the assembly sequence must maintain a balanced number of operations. An

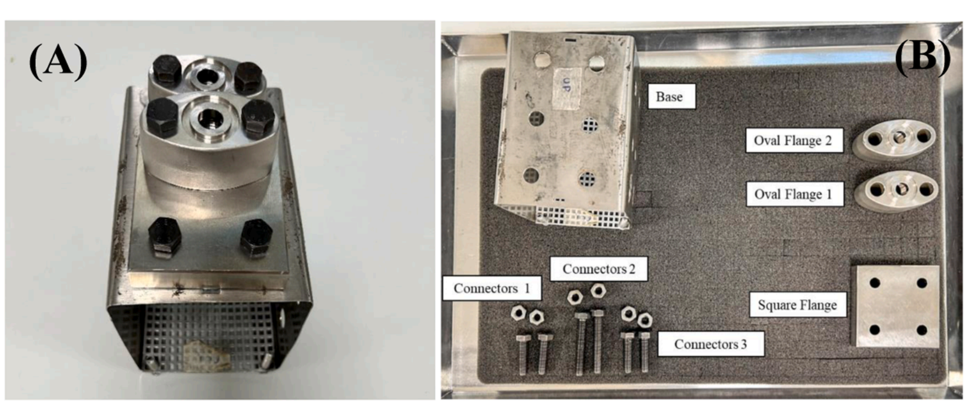


Fig. 8. Mechanical equipment used in the case study- (A) Assembled mechanical equipment; (B) Constituent parts with their identifiers.

Table 4
List of the mechanical equipment parts with their relative identifiers.

Component	Identifiers
Base	B
Square Flange	SF
Oval Flange 1	OF1
Oval Flange 2	OF2
Connectors 1	C1
Connectors 2	C2
Connectors 3	C3

excessive number of operations requires detailed, complex instructions that can compromise practicality. Conversely, overly granular segmentation leads to an increased number of instructions, which may disrupt the continuity of the assembly process. Following Stork and Schubö (2010) methodology, each task is divided into two subtasks: the *commissioning subtask*, involving all operations related to component handling, and the *joining subtask*, which covers positioning and securing the component to a subassembly.

The formalization of task subdivisions is detailed in Fig. 9, where the initial subdivision outlines a set of task-level elements. For each element, one or two subtasks are identified based on the component's characteristics. Specifically, a single commissioning subtask is assigned to a generic base component (the component on which all others are mounted). For all other components, two distinct subtasks are generated.

Table 5 provides an example of a breakdown for the assembly of mechanical equipment (see Fig. 8A and Fig. 8B).

5.2. Step 2: information requirements identifications

The second step in applying the method involves constructing a matrix to identify and select the AFs that serve as information requirements for each subtask in the assembly process. The columns of the matrix represent the subtasks, while the rows list the AFs.

The process begins with a thorough analysis of the decision-making required for each subtask, followed by the selection of relevant AFs for inclusion in the matrix.

It is essential that this selection is tailored to the operator's experience level to ensure that the information provided is aligned with their existing expertise. In this consideration, the matrix is constructed based on the operator's familiarity and proficiency with the tasks, with the aim of including only the most relevant and essential AFs for each subtask. This approach ensures that novice operators can receive detailed guidance, while experienced operators are provided with concise and relevant information.

As an example, Table 6 identifies the AFs that are information requirements for the subtask 'Square Flange (SF) Joining '. Fig. 10 provides an overview of the information requirements for all the subtasks of the mechanical equipment assembly process (see Fig. 8A and Fig. 8B), specifically tailored for a novice operator.

5.3. Step 3: mapping of information requirements to instruction formats

The third step of the methodology addresses the different formats through which the instruction content can be effectively conveyed, by establishing a correlation between the information requirements of each subtask and the instruction formats delineated in Section 3.1.

To facilitate the selection process, a matrix relating to each subtask can be created (see Fig. 11). The columns list the individual instruction formats, together with a collection of multimedia formats chosen by the designer. The rows, conversely, list the subset of AFs assessed as information requirements for the examined subtask. It is crucial to acknowledge that while the analysis permits the consideration of all possible instruction formats for comprehensiveness, real-world scenarios often impose constraints based on the instruction delivery tools at the designer's disposal. For example, if the designer has access only to a monitor, the capability is limited to delivering monitor-based instructions. Consequently, formats such as static and dynamic 3D models, which require AR technologies, must be excluded from the column of the matrix.

In detail, Step 3 can be subdivided into three sequential sub-steps, as shown below using the square flange joining subtask as an illustrative example (see Fig. 11):

1) *Matrix creation:* in the first sub-step, the matrix that links the information requirements with the instruction formats is constructed. This begins with the matrix defined in step 2, from which the list of AFs is filtered to include only those assessed as information requirements for the specific subtask being analysed. Fig. 11 exemplifies this sub-step for the SF joining subtask.
2) *Association between information requirements and instruction formats:* relying on Table 2 as a reference, the information requirements that each format can represent are determined. Fig. 12 offers a practical demonstration of this process applied to the square flange joining subtask.
3) *Definition of multimedia formats:* utilizing the matrix resulting from the previous stage as a basis, suitable multimedia formats can be

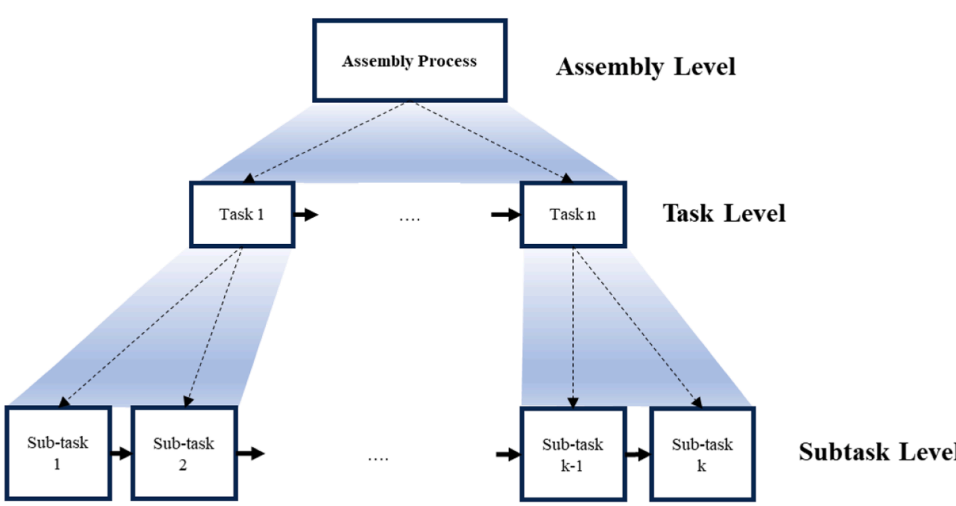


Fig. 9. Formal representation of the process subdivision.

Table 5
Practical example of the subdivision of a mechanical equipment assembly process.

Assembly Level	Task Level	Subtask Level	Description
Mechanical Equipment Assembly	**Base Fixturing**	*Base Commissioning*	The base must be fixed on the workbench by means of a clamp.
	Square Flange Assembly	*Square Flange Commissioning*	The square flange must be identified and grasped by the feeding tray.
		Square Flange Joining	The square flange must first be positioned over the base and then secured by means of screws and nuts (see Connectors 1 in Fig. 8B).
	Oval Flange 1 Assembly	*Oval Flange 1 Commissioning*	The oval flange 1 must be identified and grasped by the feeding tray.
		Oval Flange 1 Joining	The oval flange 1 must first be positioned over the square flange and then secured by means of screws and nuts (see Connectors 2 in Fig. 8B).
	Oval Flange 2 Assembly	*Oval Flange 2 Commissioning*	The oval flange 2 must be identified and grasped by the feeding tray.
		Oval Flange 2 Joining	The oval flange 2 must first be placed on the free holes in the base and then secured with nuts and screws (see Connectors 3 in Fig. 8B).

Table 6
Analysis of information requirements for the Square Flange Joining in Fig. 10.

Information Requirements	Reasoning
Feeding	Operator needs to identify the correct connector pair for assembly.
Involved form feature type	The operator without previous experience is unable to choose which holes in the base to place the square flange on. Consequently, the forms involved in the connection are an information requirement.
Insertion position	The square flange, equipped with through holes, permits two potential orientations upon the base. An operator lacking prior experience may be unable to discern the appropriate orientation to select, thus necessitating explicit instruction regarding the insertion position.
Final position	The operator must be instructed regarding the orientation and the appropriate locations for inserting the required bolts.
Insertion path	The insertion path represents the tightening of the screw. Considering the inexperience of the operator, the insertion path is considered an information requirement.

defined. The combination of individual formats into a multimedia has two purposes: (i) it enables the merging of formats that, on their own, cannot fulfil all information requirements, exploiting their complementary representational capabilities; and (ii) information elements can be distributed to the instruction format that most efficiently represents them. In this way, the processing efficiency as compared to using each format separately can be enhanced.

5.4. Step 4: instruction formats assessment and selection

In step 4, for each subtask, the instruction formats are positioned within the synthesis map (see Fig. 7). Instruction formats that provide the required representational capability and optimise the efficiency of information presentation (i.e., that fall into the 'Representative and Efficient' quadrant) can be finally selected.

In detail, Step 4 can be divided into two sequential sub-steps:

1) *Instruction formats assessment:* each instruction format is evaluated according to the two parameters described in Section 4. Using the matrices presented in Fig. 13 as a reference, the $RA_{i,j}$ value is determined using Eq. (1). For each instruction format, the value of $ReAF_{i,j}$ is ascertained by counting the number of elements within the rows. On the other hand, the value of TR_j corresponds to the total number of AFs considered as information requirements for the task under consideration. This process is repeated for all formats listed in the columns of the matrix in Fig. 13. The $PE_{i,j}$ value is derived by first assigning a $PT_{i,j}$ value to each format. The $PT_{i,j}$ value is established by considering the information requirements of the specific subtask. According to Eq. (2), the $PT_{i,j}$ value is then divided by the estimated time required for conducting the analysed subtask. Table 7 reports the assessments, along with the associated considerations, of the

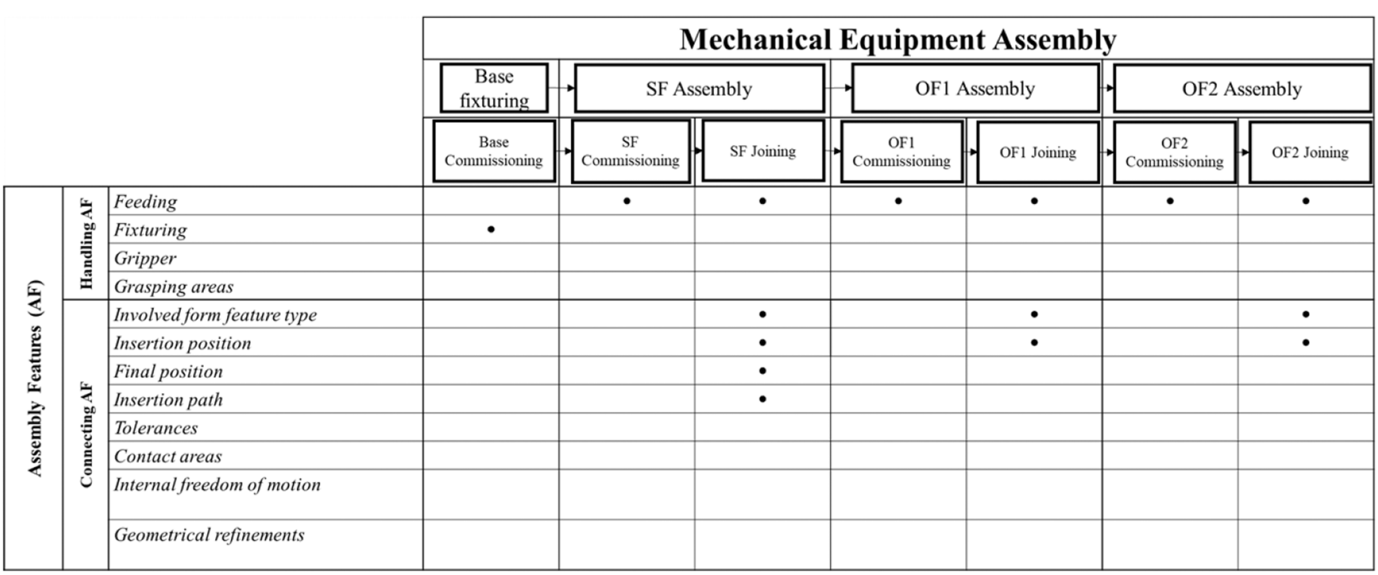


Fig. 10. Practical example of information requirements definition for the mechanical equipment assembly process. Symbol legend: SF = Square Flange; OF1 = Oval Flange 1; OF2 = Oval Flange 2.

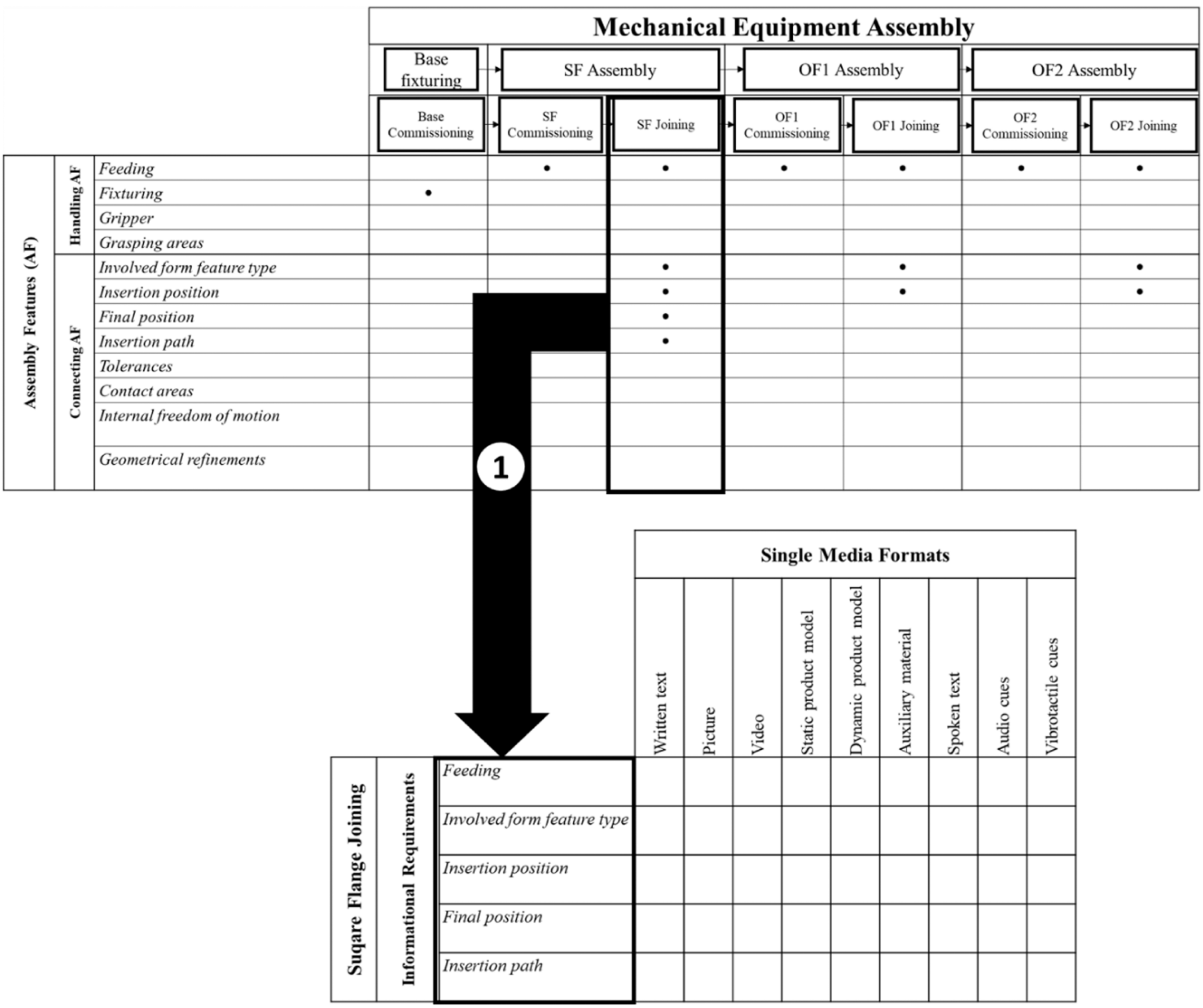


Fig. 11. Exemplification of the sub-step 1 for the square flange joining task.

Assembly Features			Single Media Formats								
			Written text	Picture	Video	Static product model	Dynamic product model	Auxiliary materials	Spoken text	Audio cues	Vibrotactile cues
Assembly Features	Handling AF	*Feeding*	•	•	•	•	•	•	•	•	•
		Fixturing	•		•		•		•		
		Gripper	•		•		•		•		
		Grasping areas	•		•		•	•	•		
	Connecting AF	*Involved form feature type*	•	•	•	•	•	•	•	•	•
		Insertion position	•	•	•	•	•		•		
		Final position	•	•	•	•	•		•		
		Insertion path	•		•		•	•	•		
		Tolerances	•						•		
		Contact areas	•		•		•	•	•		
		Internal freedom of motion	•		•		•	•	•		
		Geometric refinements	•	•	•	•	•		•		

2

Suqare Flange Joining			Single Media Formats								
			Written text	Picture	Video	Static product model	Dynamic product model	Auxiliary material	Spoken text	Audio cues	Vibrotactile cues
Suqare Flange Joining	Informational Requirements	*Feeding*	•	•	•	•	•	•	•	•	•
		Involved form feature type	•	•	•	•	•	•	•	•	•
		Insertion position	•	•	•	•	•		•		
		Final position	•	•	•	•	•		•		
		Insertion path	•		•		•	•	•		

Fig. 12. Exemplification of the sub-step 2 for the square flange joining subtask.

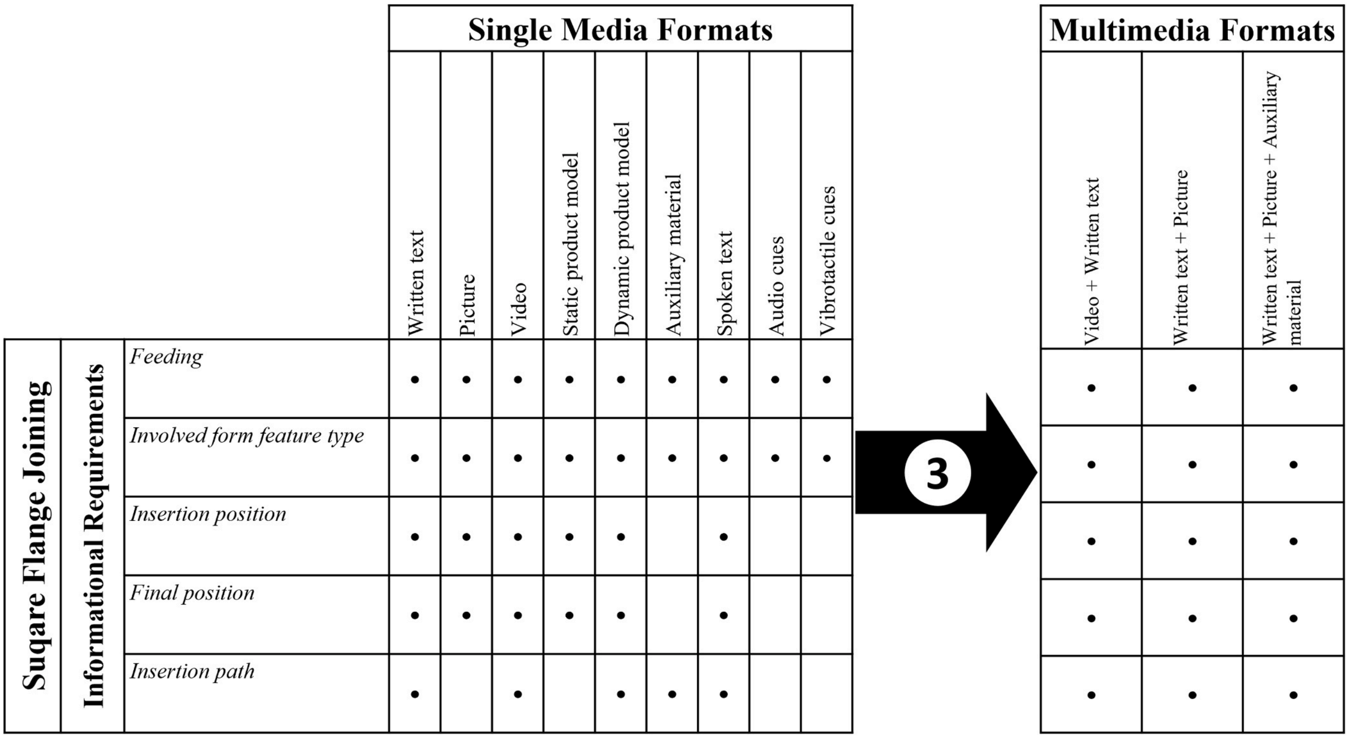


Fig. 13. Exemplification of the sub-step 3 for the square flange joining subtask.

Table 7
$RA_{1,3}$ and $PE_{1,3}$ assessment for the written text format for the SF Joining in Fig. 7.

Written Text		
$RA_{1,3}$= 1	$ReAF_{1,3} = 5$	Written text can represent all five AFs which are information requirements for the third task.
	$TR_3 = 5$	Five AFs have been assessed as information requirements to represent the third subtask.
$PE_{1,3}$= 2	$PT_{1,3}$= 20 s	The verbal representation of the mutual position of the base and the square flange, their orientation and finally the screwing generate a long and complex textual instruction. The processing time for the operator is estimated at 20 s.
	TT_3= 10 s	The estimated average duration of the square flange joining subtask is 10 s.

written text formats related to the subtask of joining the square flange. The written text corresponds to the first format ($i = 1$), and the assessment focuses on the joining of the square flange (j = 3). Comprehensive results for all instruction formats can be found in Fig. 14.

2) *Instruction format mapping and selection:* based on the results of the assessment conducted in the previous step, the instruction formats are allocated on the synthesis map. For the exemplificative case study, the position of single and multimedia formats assessments is reported in Fig. 15. Based on this map, the instruction designer can evaluate and select the format that achieves a balanced mix of representational ability and efficiency in presenting instruction information. In the case of the joining subtask involving the square flange, the ideal format is a multimedia presentation of written text and picture. If multiple formats appear in the optimal quadrant, the designer should choose based on the lowest $PE_{i,j}$ values, indicating greater efficiency for equivalent representational capability. If these values are equal, selection may be refined based on the practical difficulty of creating the instructions.

6. S-AID effectiveness preliminary verification

6.1. Experimental investigation

The proposed experimentation aims to preliminary evaluate the effectiveness of instructions designed using the *S-AID* method. The assembly process of the mechanical component used to describe the *S-AID* method was used as a case study (see Fig. 8). This process was selected as it requires a sufficient degree of dexterity to screw the components together. In addition, a relevant level of attention is required for their positioning. Multiple assembly positions can be selected and components can be positioned in different orientations, making the use of instructions significant to the success of the assembly process. The standard completion time for the assembly of the mechanical equipment is around 220 s.

The set of instruction created with *S-AID* method were evaluated against three traditional instruction types—text, video, and picture—all displayed on a monitor. A sample of 40 students (15 males and 25 females), aged 23–27 and inexperienced in the assembly task, participated in a between-groups design study. Each subject, randomly assigned to one of the instruction types, performed the assembly process once to isolate the instruction type's factor impact and prevent the learning effects arising from task repetition.

The effectiveness of the *S-AID* versus traditional instructions was measured using the following metrics:

- *Assembly completion time:* the time required to complete the task was recorded. This variable is a dependable proxy for the operator's level of understanding of the assembly sequence and of the timeliness of instructions.
- *Number of process failures:* during the task, the number of process failures was recorded. Common examples of process failures are incorrect selection, incorrect positioning (both assembly position

	Single Media Formats									Multimedia Formats		
	Written text	Picture	Video	Static product model	Dynamic product model	Auxiliary material	Spoken text	Audio cues	Vibrotactile cues	Video + Written text	Written text + Picture	Written text + Picture + Auxiliary material
$RA_{i,j}$	1	0.8	1	0.8	1	0.6	1	0.4	0.4	1	1	1
$PE_{i,j}$	2	0.5	1.5	0.5	1.5	0.25	2	0.25	0.25	2	0.75	1.25

Fig. 14. Assessment of instruction formats for the SF joining in Fig. 7.

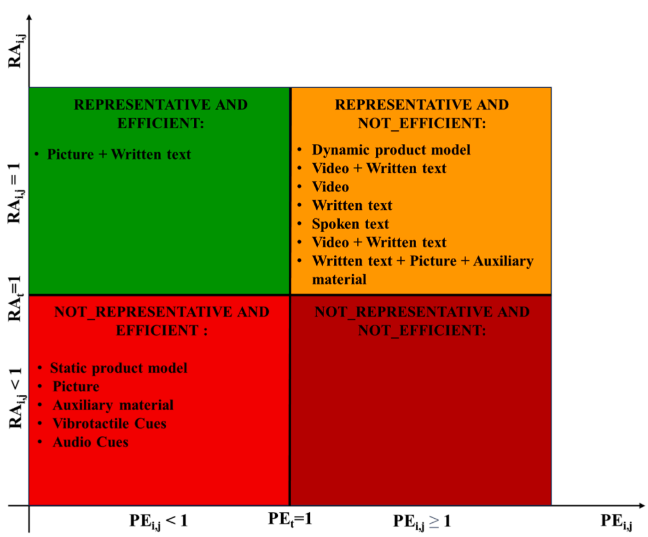


Fig. 15. Synthesis map of instruction formats for the SF joining in Fig. 8.

and orientation) and incorrect assembly of the component. This performance metric allows us to understand the efficacy of the instructions.

- *Perceived cognitive load score:* cognitive load refers to the amount of cognitive resources needed to process information and complete a task. In this study, the revised version of the NASA Task Load Index (NASA-TLX) proposed by Gerjets et al. (2006) was employed to measure perceived cognitive load. This scale conceptualizes cognitive load as comprising three distinct items: (i) *Task Demand*, which measures how much mental or physical activity is required for the task, reflecting the intrinsic difficulty of the task; (ii) *Effort*, which assesses how much effort is put into understanding the instructions needed to perform the task, associated the cognitive resources used to develop schemas of task performance; and (iii) *Navigation Demand*, which gauges the effort required to navigate through the instructions in search of the needed information, associated with extraneous load due to potentially suboptimal instruction design. Each of these items is rated on a scale from 0 to 100. Aggregate perceived cognitive load scores were calculated using an equal weighting across these dimensions to signify their equal relevance.

6.2. Instructions design set-up

The set-up phase of the experiment involved the creation of the instruction sets to be compared during the experiment.

First, the *S-AID* method was applied to the mechanical equipment. The result of Step 1 and Step 2 of the method are illustrated together in Fig. 10. The evaluations were conducted on the premise of the inexperience of the instruction recipients. Subsequently, from the formats shown in Table 3, those that can be represented through the user interface available in the assembly station, a monitor, were selected (see the instruction formats in the first column of Fig. 16). Each format was evaluated on *RA* and *PE* dimensions on each subtask. The results of Step 3 and Step 4 of the method are summarised in Fig. 16. An example of an *S-AID*-developed instruction is shown in Fig. 18.

The three following alternative instruction sets, reflecting “traditional” method of instruction design, were developed:

- *Text-based instruction*: for each subtask, the actions to be performed were described verbally in a minimal manner.
- *Video-based instruction*: a video was created for the assembly of each component.
- *Picture-based instruction:* the format comprises a set of successive illustrations that depict the assembly procedure of the mechanical equipment, with each component represented by a diagram (for example see Fig. 17).

It is notable that the evaluation of spoken text was not included in the list of traditional formats. This is due to the fact that spoken text has comparable characteristics to written text, making its evaluation potentially redundant. Furthermore, in the context of this assembly, the length of the text to be provided in audio mode would have accentuated its inefficiency in terms of timeliness. In particular, as outlined in Section 3.2, it is inefficient to quickly locate and rehearse important information in stand-alone spoken instructions (Dunham et al., 2020).

6.3. Experimental procedure

The experiment involved assigning each participant a specific type of

Instruction Formats		Base fixturing: Base Commissioning	SF Assembly: SF Commissioning	SF Assembly: SF Joining	OF1 Assembly: OF1 Commissioning	OF1 Assembly: OF1 Joining	OF2 Assembly: OF2 Commissioning	OF2 Assembly: OF2 Joining
Written text	RA	1	1	1	1	1	1	1
	PE	2	1.5	2	1.5	1,25	1.5	1,25
Picture	RA	0	1	0.8	1	1	1	1
	PE	0	1.25	0.5	1.25	0,75	1.25	0,5
Video	RA	1	1	1	1	1	1	1
	PE	1	1,25	1.5	1,25	1.25	1.25	1.25
Spoken text	RA	1	1	1	1	1	1	1
	PE	2.25	2	2	2	1,5	2	1,5
Written text + Picture + Auxiliary material	RA	1	1	1	1	1	1	1
	PE	1.25	0.75	1.25	0.75	1.25	0.75	1.25
Picture + Auxiliary material	RA	0	1	1	1	1	1	1
	PE	0	0.5	0.75	0.75	0.75	0.5	0.75
Picture + Spoken text	RA	1	1	1	1	1	1	1
	PE	1.75	0.75	1	0.75	1.25	0.75	1.25
Picture + Written text	RA	1	1	1	1	1	1	1
	PE	1.25	0.75	0.75	0.75	0.75	0.75	0.75
Video + Audio text	RA	1	1	1	1	1	1	1
	PE	1	1.25	1.25	1.25	1.25	1.25	1.25
Video + Auxiliary material	RA	1	1	1	1	1	1	1
	PE	1.5	1.25	1.5	1.25	1.25	1.25	1.25
Selected instruction format		*Video*	*Picture + Auxiliary Material*	*Picture + Written text*	*Picture + Auxiliary Material*	*Picture + Written text*	*Picture + Auxiliary Material*	*Picture*

Fig. 16. Illustration of the result of the application of the S-AID method to the mechanical equipment assembly process.

instructions to perform a one-time assembly of the mechanical equipment. At the beginning of the experiment, participants were informed of the study's objectives and the identifiers of the mechanical equipment components. This ensured that they could quickly associate the component names mentioned in the textual instructions with the correct parts in the work area. Subsequently, participants were trained on how to interact with the monitor interface, including how to scroll through and review instructions. This training ensured that each participant could navigate the instructions with ease. Once participants confirmed their comfort with the interface, they were given the opportunity to commence the trial by giving a start signal, thereby synchronizing the trial's start with the experimenters' collection of completion times. The operations executed by the participants during the assembly of the mechanical equipment are detailed in Table 5. During assembly execution, any process failures were recorded. Assembly was considered complete when the participant placed the assembled product back into the feeding tray. Following the completion of the assembly, participants were invited to complete the modified NASA-TLX to assess their perceived cognitive load. Finally, participants were also asked to provide unstructured feedback regarding the completeness and immediacy of the instructions.

6.4. Results

This section details the results of the comparison of instructions created using the *S-AID* methodology with "traditional" single-format instructions. The analysis focuses specifically on the response variables of assembly completion time, process failures, and perceived cognitive load. Perceived cognitive load analysis also involves the dimensions of task demand, effort and navigation demand.

The normality of the distributions for each response variable under different instruction types of conditions was assessed using the Shapiro-Wilk test. The test results indicate that the normality hypothesis cannot be rejected for the distributions of assembly completion time, aggregate modified NASA TLX perceived cognitive load score and its dimensions. Consequently, one-way ANOVA tests were performed to analyse the influence of the instruction type factor on each of the response variable. To investigate the presence of significant differences between the groups individuated by instruction types, a post-hoc analysis with pairwise *t*-test and Bonferroni correction was performed for each response variable (Lee and Lee, 2018).

The normality hypothesis was rejected for process failures distributions. Consequently, to analyse the influence of the instruction type factor, the non-parametric test Kruskal-Wallis was adopted. Subsequently, to analyse the differences between the groups, a Wilcoxon rank-sum post-hoc test with Bonferroni correction was performed.

6.4.1. Assembly completion time

Fig. 19 provides a comparative analysis of assembly completion times across different instruction sets, with one-way ANOVA results indicating a significant influence of instruction type on assembly completion time ($p < 0.001$). Analysing the differences between the different types of instruction, emerges that assembly processes conducted with *S-AID-based* instructions present the lowest median. This finding is further supported by post-hoc analyses, which identify significant differences between *S-AID-based* instructions and both *text-*

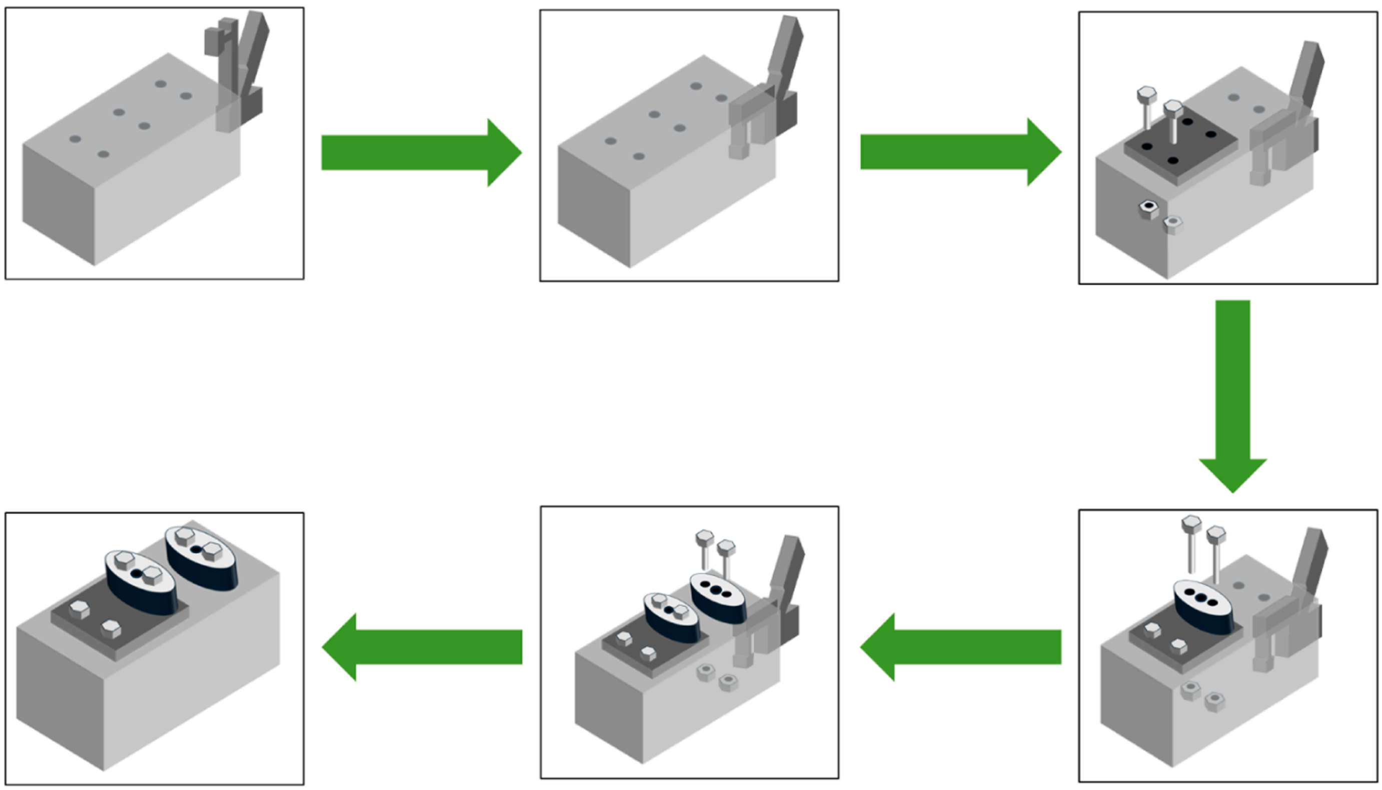


Fig. 17. Set of picture-based instructions used in the experiment for the assembly of the mechanical equipment. This figure presents the complete set of picture-based instructions as displayed together for illustrative purposes. During the experiment, the illustrated pictures were provided to participants sequentially, one at a time and without the arrows.

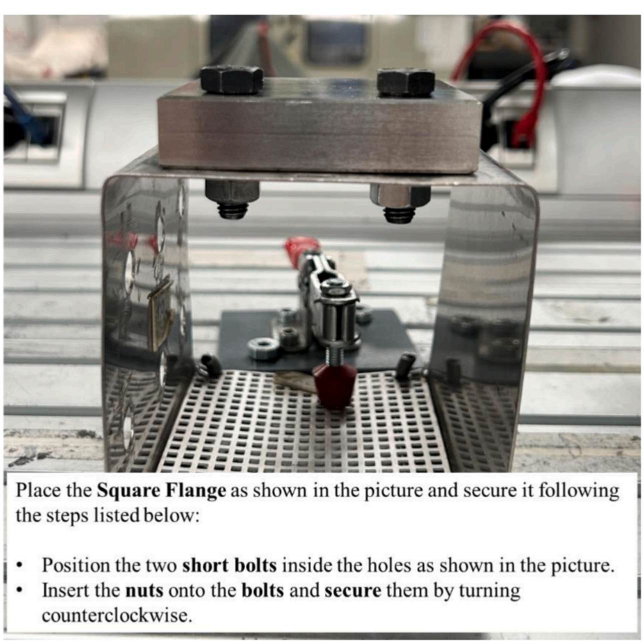


Fig. 18. Example of written text + picture multimedia instruction provided for the square flange assembly subtask using the S-AID method.

based and *video-based* instructions, as detailed in Table 8. The longer duration of the assembly processes with *video-based* instructions can be explained by participants' tendency to watch the entire video before starting the task, leading to time inefficiency, especially for simpler operations. This contrasts with the effectiveness of *S-AID* in producing immediate instructions tailored to specific task characteristics. *Text-based* instructions proved to be the least time-efficient. Significantly longer assembly completion times are observed with *text-based* instructions compared to all other instruction types ($p < 0.001$ for all comparisons). This effect arises from the need to carefully read and align textual instructions with real-world conditions, a process that is not only time-consuming but also heavily dependent on individual capabilities. Moreover, correcting process failures in tasks guided by textual instructions typically necessitates detailed re-reading, which significantly extends the duration of the task. Finally, no significant difference resulted between *S-AID-based* and *picture-based* instructions ($p = 0.355$).

6.4.2. *Process failures*

Fig. 20 displays the distribution of the number of process failures that occurred during assembly execution with the four different instruction sets.

The Kruskal-Wallis test demonstrates that the instruction set significantly influences the number of process failures ($p < 0.001$). A further analysis of the differences between the types of instructions revealed that the highest number of failures occurred when the task was executed with *picture-based* instructions (see Fig. 20). The Wilcoxon rank-sum post-hoc test with Bonferroni correction indicates that this effect is statistically significant with respect to both *S-AID-based* and *video-based* instructions (see Table 9). This variation primarily manifests in the operations of base fixturing and bolts tightening. In these activities, the participant must infer the required manual actions, as they are not explicitly deducible from the pictures of the component. On the contrary, *S-AID-based* and *video-based* instructions explicitly provides this information, thus eliminating the operator's need to infer actions and reducing uncertainty. Furthermore, trials conducted with *S-AID-based* instructions exhibit fewer process failures compared to those using *video-based* and *text-based* instruction. However, according to the Wilcoxon-rank sum test, this effect is not statistically significant (see

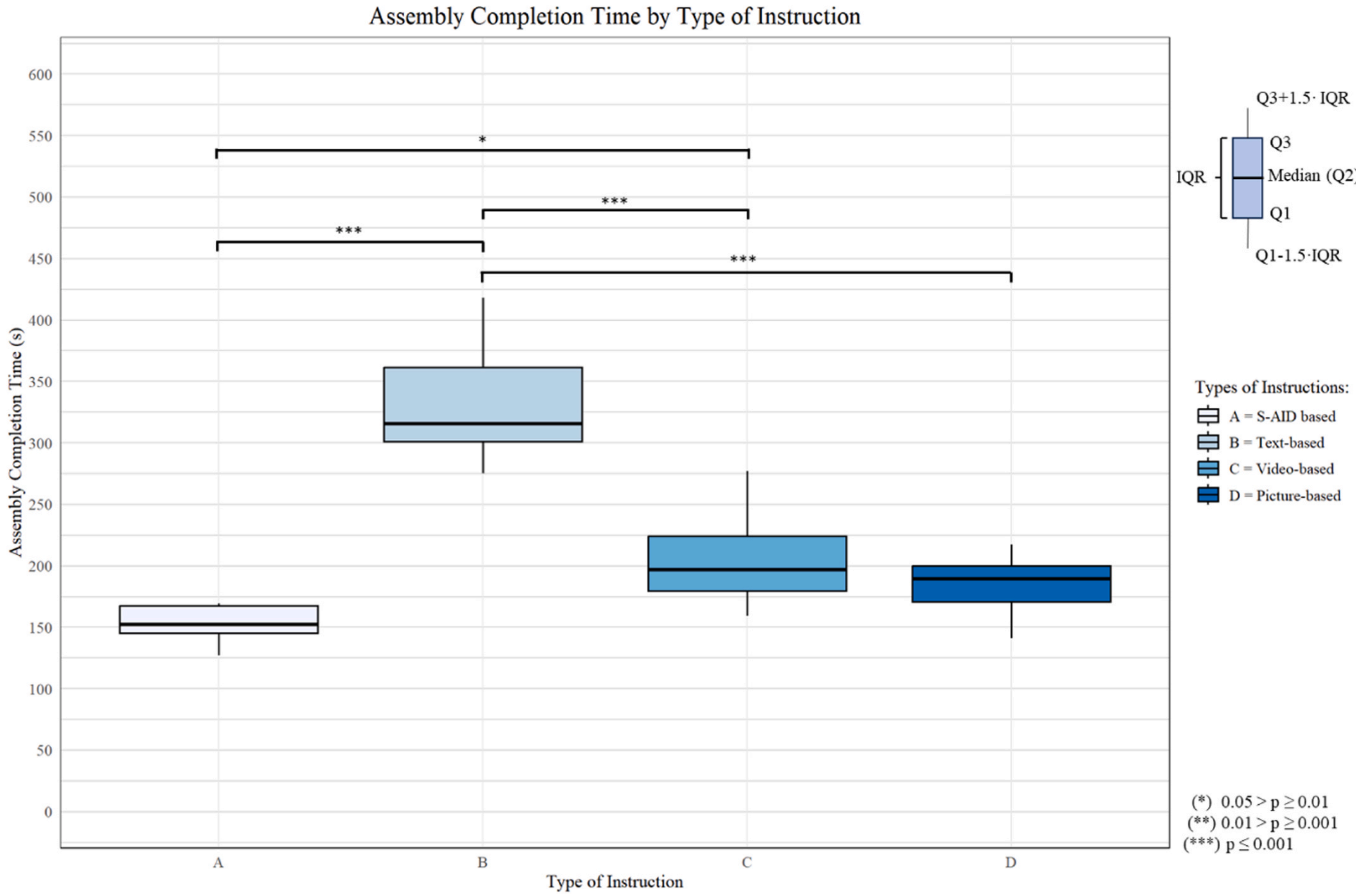


Fig. 19. Boxplot comparison of assembly completion times between different types of instruction.

Table 8
Results of statistical analyses for assembly completion times.

Post-hoc analysis (pairwise *t*-test with Bonferroni correction)	
Comparison	*p-value adjusted*
S-AID based vs. Text-based	$p < 0.001$
S-AID based vs. Video-based	$p = 0.012$
S-AID based vs. Picture-based	$p = 0.355$
Text-based vs. Video-based	$p < 0.001$
Text based vs. Picture-based	$p < 0.001$
Video-based vs. Picture-based	$p > 0.999$

Table 9).

6.4.3. Perceived cognitive load score

Fig. 21 displays a graphical comparison of the distributions of the scores of each dimension of the modified NASA-TLX, as well as its aggregate score, across the different instruction types.

One-way ANOVA analysis showed a statistically significant impact of instruction type on Task Demand scores ($p = 0.0051$). Post-hoc analyses revealed that *text-based* instructions were associated with higher perceived task difficulty than *S-AID*-based instructions ($p = 0.012$), likely due to their complex nature which requires extensive semantic processing before being completely understood.

For the Effort dimension, which assesses commitment to learning the task, the results of the one-way ANOVA showed that the type of instruction significantly affected the scores reported ($p = 0.0098$). A significantly higher effort score was observed for *picture-based* instructions, with a statistically significant difference reported with respect to *video-based* instructions ($p = 0.006$). On the contrary, *S-AID-based* instructions were generally perceived as less demanding than both *text* and *picture-based* instructions. However, according to the pairwise *t*-test with Bonferroni correction, these differences were not statistically significant (see Table 10).

The Navigation Demand scores, which reflects the effort to find the required information in the instructions, were significantly influenced by the type of instruction ($p < 0.001$). *S-AID-based* instructions were deemed clear and immediate, significantly reducing navigation demand compared to *text-based* instructions ($p < 0.001$). Although *S-AID-based* instructions showed generally lower perceived navigation demand, these differences were not significant compared to *picture* and *video-based* instructions (see Table 10).

As evidenced by the distribution of the aggregate scores of the modified NASA-TLX, overall perceived cognitive load was generally higher for text-based instruction. Significant differences were found compared to *S-AID-based* and *video-based* instructions ($p < 0.001$ for both). *Video-based* instructions reported the lowest perceived cognitive load scores. The observed difference was found to be significant with respect to *picture-based* instructions ($p = 0.005$), but not with regard to *S-AID-based* and *video-based* instructions. The minor differences in perceived cognitive load between *video-based* instructions and both *S-AID-based* and *picture-based* instructions are primarily due to variations in reported Effort dimension scores. This can be seen as a positive aspect, demonstrating enhanced engagement and involvement in learning tasks with *S-AID* and *picture-based* instructions.

7. Discussion and conclusions

This paper introduced the *Streamlined Assembly Instruction Design* (*S-AID*) *method*, a systematic approach tailored to enhance the instruction

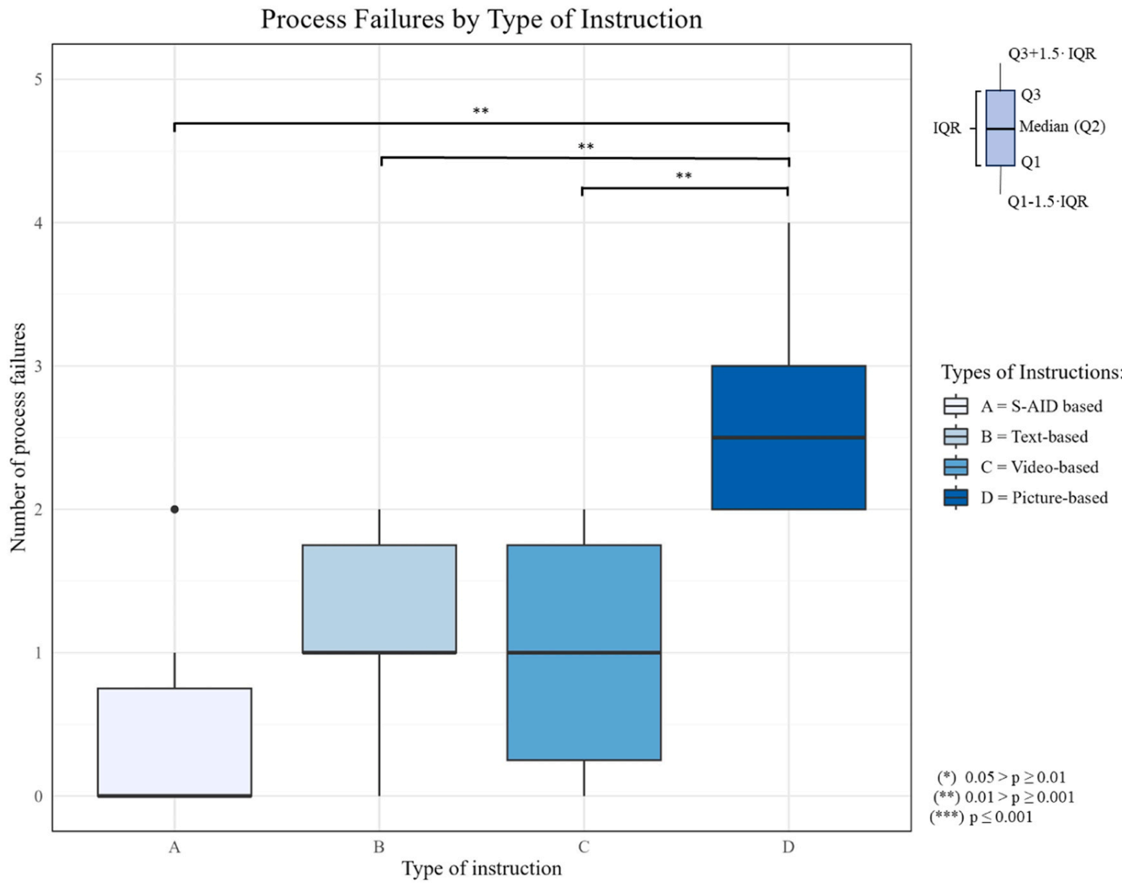


Fig. 20. Boxplot comparison of the number of process failures between different types of instruction.

Table 9
Results of statistical analyses for process failures.

Post-hoc analysis (Wilcox rank-sum test with Bonferroni correction)	
Comparison	*p-value adjusted*
S-AID based vs. Text-based	$p = 0.249$
S-AID based vs. Video-based	$p = 0.548$
S-AID based vs. Picture-based	$p = 0.0013$
Text-based vs. Video-based	$p > 0.999$
Text based vs. Picture-based	$p = 0.0054$
Video-based vs. Picture-based	$p = 0.0056$

design process in assembly operations. Central to *S-AID* is the understanding of the cognitive mechanisms that influence decision-making during assembly, allowing for a detailed analysis of the information requirements essential for effective task execution. By proposing a taxonomy of *Assembly Features (AFs)*, *S-AID* equips designers with a robust tool to identify and categorize necessary information, aiding in the precise crafting of instructions.

The differentiation and classification of instruction formats also stand as a specific contribution introduced by this paper. This classification helps designers align the instruction format with the specific characteristics and requirements of each assembly task. Furthermore, *S-AID* introduces two critical assessment variables— *Representational Ability* and *Processing Efficiency* —that guide the selection of the most appropriate format to optimize both correct execution understanding and efficiency of information presentation.

The S-AID method exhibits robust adaptability to various technological constraints, operator experience levels, and assembly processes. Specifically, the method allows for the exclusion of technologically incompatible instruction formats and adjusts information requirements based on the operator's experience and learning style, thus ensuring broad applicability and scalability across different operative scenarios. The initial step of subdividing the assembly process into subtasks follows the framework proposed by Stork and Schubö (2010), which is generalisable to the majority of assembly processes. Furthermore, the information requirement identification step is based on exhaustive taxonomy of AFs derived from the Design for Assembly (DFA) literature (Van Holland and Willem, 2000), ensuring its applicability for a wide range of assemblies. This conceptual foundation can be considered as a preliminary indication that, although the applicability of the method has been showed for a single case study, it has the potential to be extended to most of assembly processes and operational scenarios.

To verify its effectiveness, *S-AID* method was compared with traditional text, video and picture-based instructions in a case study. The results showed that *S-AID*-based instructions reduces assembly completion times and process failures while maintaining an acceptable level of perceived cognitive load compared to other instruction types. Overall, these outcomes demonstrated that *S-AID-based* instructions achieve an optimal balance between exhaustiveness, as evidenced by fewer process failures, and efficiency, as shown by reduced assembly completion times.

By providing a structured framework for assembly instruction design, *S-AID* approach introduces a systematic way to match instruction formats with the specific information requirements of different tasks. The comprehensive taxonomy of AFs and the assessment variables for representational ability and immediacy offer valuable insights that can guide further research and improvements in instruction design. This could lead to improvements in productivity and skill acquisition across various industries.

Despite these strengths, it is important to acknowledge *S-AID* method limitations. Firstly, the process of identifying cognitive processes to

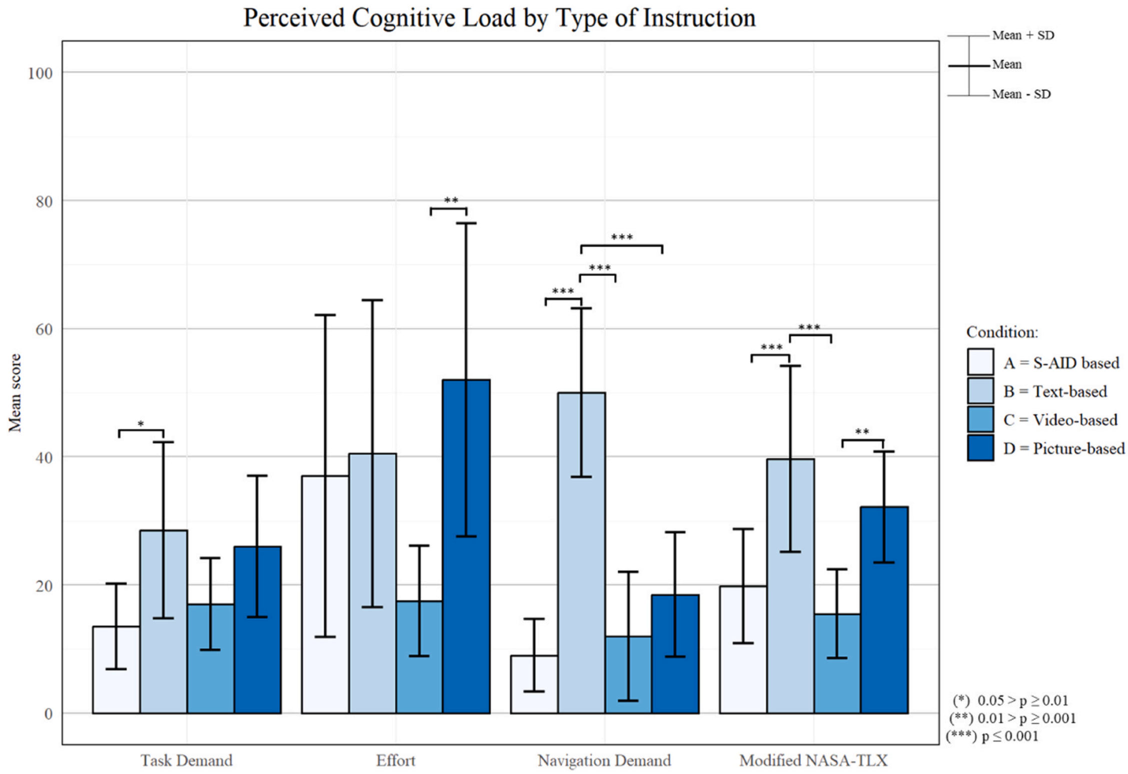


Fig. 21. mean scores with standard deviation of modified NASA-TLX and its dimensions for the different instruction types.

Table 10
Statistical analyses of the scores of the modified NASA TLX and its dimensions.

Response Variable	Post hoc: S-AID-based vs.		
	Text-based	*Video-based*	*Picture-based*
Task Demand	$p = 0.012$	$p > 0.999$	$p = 0.052$
Effort	$p > 0.999$	$p = 0.308$	$p = 0.779$
Navigation Demand	$p < 0.001$	$p > 0.999$	$p = 0.240$
Modified NASA-TLX	$p < 0.001$	$p > 0.999$	$p = 0.061$

determine information requirements can be complex and resource-intensive, which may detract from its utility in smaller-scale or less resource-abundant environments. Secondly, while *S-AID* is specifically designed for assembly processes, its applicability to other contexts remains unexplored. Finally, the effectiveness of the *S-AID* method was preliminary verified on a single assembly process, with the product being sufficiently representative of typical assembly challenges. However, to comprehensively evaluate its efficacy, further studies are required. These studies should aim to operationally validate the *S-AID* method across diverse real-world assembly contexts. Future research should also focus on simplifying and streamlining the *S-AID* method further, refining the assessment of instruction formats to better gauge their representational abilities and the perceived cognitive loads they impose. Digitizing and automating the process could significantly reduce the application burden, making the method more efficient and accessible. By incorporating these technological advancements, the S-AID method could adapt to various assembly scenarios more effectively, minimize human process failures, and ensure consistent quality in instruction delivery.

CRediT authorship contribution statement

Federico Barravecchia: Methodology, Data curation. **Mirco Bartolomei:** Methodology, Formal analysis, Conceptualization. **Davide Maria Gatta:** Project administration. **Luca Mastrogiacomo:** Supervision, Methodology. **Fiorenzo Franceschini:** Supervision, Methodology, Conceptualization.

Ethical approval

The authors respect the Ethical Guidelines of the Journal.

Author statement

NA

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Mirco Bartolomei reports financial support was provided by Ministero dell'Università MUR. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This publication is part of the project PNRR-NGEU which has received funding from the MUR – DM 352/2022 and a collaboration between Politecnico di Torino and Leonardo S.p.A. We acknowledge the support of Leonardo S.p.A. for this study, specifically to Francesco Acerra and Davide Cannizzaro, for their crucial insights and expert guidance throughout the collaboration. The authors and Leonardo S.p.A jointly own the intellectual property of this work.

Authors contributions

The authors have provided an equal contribution to the drafting of the paper.

Data availability

No data was used for the research described in the article.

References

Alessa, Faisal M., Alhaag, Mohammed H., Al-harkan, Ibrahim M., Ramadan, Mohamed Z., Alqahtani, Fahad M., 2023. A neurophysiological evaluation of cognitive load during augmented reality interactions in various industrial maintenance and assembly tasks. Sensors 23 (18). https://doi.org/10.3390/s23187698.

Arbeláez, J.C., Viganò, Roberto, Osorio-Gómez, Gilberto, 2019. Haptic augmented reality (HapticAR) for assembly guidance. Int. J. Interact. Des. Manuf. 13 (2), 673–687. https://doi.org/10.1007/s12008-019-00532-3.

Arguel, Amaël, Jamet, Eric, 2009. Using video and static pictures to improve learning of procedural contents. Comput. Hum. Behav. 25 (2), 354–359. https://doi.org/10.1016/j.chb.2008.12.014.

Atkinson, R.C., Shiffrin, R.M., 1968. Human memory: a proposed system and its control processes. Psychol. Learn. Motiv. - Adv. Res. Theory 2 (C), 89–195. https://doi.org/10.1016/S0079-7421(08)60422-3.

Ayres, Paul, Sweller, John, 2014. The Split-Attention Principle in Multimedia Learning. The Cambridge Handbook of Multimedia Learning, Second Edition. Cambridge University Press, Cambridge.

Baddeley, A.D. 1979. *Working Memory and Reading.* edited by M. E. and B. H. Kolers Paul A. and Wrolstad. Boston, MA: Springer US.

Barravecchia, Federico, Bartolomei, Mirco, Mastrogiacomo, Luca, Franceschini, Fiorenzo, 2023. Redefining human–robot symbiosis: a bio-inspired approach to collaborative assembly. Int. J. Adv. Manuf. Technol. https://doi.org/10.1007/s00170-023-11920-1.

Bieger, George R., Glock, Marvin D., 1985. The information content of picture-text instructions. J. Exp. Educ. 53 (2), 68–76. https://doi.org/10.1080/00220973.1985.10806364.

Bieger, George R., Glock, Marvin D., 1986. Comprehending spatial and contextual information in picture-text instructions. J. Exp. Educ. 54 (4), 181–188. https://doi.org/10.1080/00220973.1986.10806418.

Boothroyd, G., Alting, L., 1992. Design for assembly and disassembly. CIRP Ann. - Manuf. Technol. 41 (2), 625–636. https://doi.org/10.1016/S0007-8506(07)63249-1.

Canales, Ryan, Jörg, Sophie, 2020. Performance is not everything: audio feedback preferred over visual feedback for grasping task in virtual reality. Proceedings of the 13th ACM SIGGRAPH Conference on Motion, Interaction and Games. https://doi.org/10.1145/3424636.3426897.

Capponi, Matteo, Gervasi, Riccardo, Mastrogiacomo, Luca, Franceschini, Fiorenzo, 2024. Assessing perceived assembly complexity in human-robot collaboration processes: a proposal based on Thurstone's law of comparative judgement. Int. J. Prod. Res. 62 (14), 5315–5335. https://doi.org/10.1080/00207543.2023.2291519.

Chandler, Paul, Sweller, John, 1991. Cognitive Load Theory And The Format Of Instruction. Cogn. Instr. 8 (4), 293–332. https://doi.org/10.1207/s1532690xci0804_2.

De Fazio, T.L., Edsall, A.C., Gustavson, R.E., Hernandez, J., Hutchins, P.M., Leung, H.W., Luby, S.C., Metzinger, R.W., Nevins, J.L., Tung, K., Whitney, D.E., 1993. A prototype of feature-based design for assembly. J. Mech. Des. Trans. ASME 115 (4), 723–734. https://doi.org/10.1115/1.2919261.

Del Missier, Fabio, Mäntylä, Timo, Hansson, Patrik, Bruine De Bruin, W.ändi, Parker, Andrew M., Nilsson, Lars G.öran, 2013. The multifold relationship between memory and decision making: an individual-differences study. J. Exp. Psychol.: Learn. Mem. Cogn. 39 (5), 1344–1364. https://doi.org/10.1037/a0032379.

Dorloh, Halimoh, Li, Kai-Way, Khaday, Samsiya, 2023. Presenting job instructions using an augmented reality device, a printed manual, and a video display for assembly and disassembly tasks: what are the differences? Appl. Sci. 13 (4). https://doi.org/10.3390/app13042186.

Dunham, Sabrina, Lee, Edward, Persky, Adam M., 2020. The psychology of following instructions and its implications. Am. J. Pharm. Educ. 84 (8), 1052–1056.

Fiorentino, Michele, Uva, Antonio E., Gattullo, Michele, Debernardis, Saverio, Monno, Giuseppe, 2014. Augmented reality on large screen for interactive maintenance instructions. Comput. Ind. 65 (2), 270–278. https://doi.org/10.1016/j.compind.2013.11.004.

Franceschini, Fiorenzo, Galetto, Maurizio, Maisano, Domenico, 2019. Designing Performance Measurement Systems: Theory and Practice of Key Performance Indicators. Springer, Cham.

Ganier, Franck, 2004. Factors affecting the processing of procedural instructions: implications for document design. IEEE Trans. Prof. Commun. 47 (1), 15–26. https://doi.org/10.1109/TPC.2004.824289.

Geng, Junhao, Song, Xinyu, Pan, Yuntao, Tang, Jianjun, Liu, Yu, Zhao, Dongping, Ma, Yongsheng, 2020. A systematic design method of adaptive augmented reality work instruction for complex industrial operations. Comput. Ind. 119, 103229. https://doi.org/10.1016/j.compind.2020.103229.

Gerjets, Peter, Scheiter, Katharina, Catrambone, Richard, 2006. Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? Learn. Instr. 16 (2), 104–121. https://doi.org/10.1016/j.learninstruc.2006.02.007.

Gervasi, Riccardo, Mastrogiacomo, Luca, Franceschini, Fiorenzo, 2020. A Conceptual Framework to Evaluate Human-Robot Collaboration. Int. J. Adv. Manuf. Technol. 108 (3), 841–865. https://doi.org/10.1007/s00170-020-05363-1.

Han, Wenkai, Schulz, Hans Jörg, 2020. Exploring Vibrotactile Cues for Interactive Guidance in Data Visualization. ACM International Conference Proceeding Series.

Hegarty, Mary, 2012. Multimedia Learning About Physical Systems. The Cambridge Handbook of Multimedia Learning. Cambridge University Press, Cambridge.

Heiser, Julie, Tversky, Barbara, 2006. Arrows in comprehending and producing mechanical diagrams. Cogn. Sci. 30 (3), 581–592. https://doi.org/10.1207/s15516709cog0000_70.

Höffler, Tim N., Leutner, Detlev, 2007. Instructional animation versus static pictures: a meta-analysis. Learn. Instr. 17 (6), 722–738. https://doi.org/10.1016/j.learninstruc.2007.09.013.

Hu, S.J., Ko, J., Weyand, L., Elmaraghy, H.A., Lien, T.K., Koren, Y., Bley, H., Chryssolouris, G., Nasr, N., Shpitalni, M., 2011. Assembly system design and operations for product variety. CIRP Ann. - Manuf. Technol. 60 (2), 715–733. https://doi.org/10.1016/j.cirp.2011.05.004.

Huk, Thomas, 2006. Who benefits from learning with 3D models? The case of spatial ability. J. Comput. Assist. Learn. 22, 392–404.

Kalyuga, Slava, Paul, Chandler, John, Sweller, 1999. Managing split-attention and redundancy in multimedia instruction. Appl. Cogn. Psychol. 13 (4), 351–371. https://doi.org/10.1002/(SICI)1099-0720(199908)13:4< 351::AID-ACP589> 3.0.CO;2-6.

Laviola, Enricoandrea, Emmanuele Uva, Antonio, Gattullo, Michele, 2024. The minimal AR authoring approach: validation in a real assembly scenario. Comput. Ind. 154. https://doi.org/10.1016/j.compind.2023.104026.

Lee, Sangseok, Lee, Dong Kyu, 2018. What is the proper way to apply the multiple comparison test? Korean J. Anesthesiol. 71 (5), 353–360. https://doi.org/10.4097/kja.d.18.00242.

Lee, Doo Young, Shin, Dong Hee, 2012. An empirical evaluation of multi-media based learning of a procedural task. Comput. Hum. Behav. 28 (3), 1072–1081. https://doi.org/10.1016/j.chb.2012.01.014.

Loch, Frieder, Quint, Fabian, Brishtel, Iuliia, 2016. Comparing Video and Augmented Reality Assistance in Manual Assembly. Proceedings - 12th International Conference on Intelligent Environments, IE 2016 147–150. https://doi.org/10.1109/IE.2016.31.

Miller, George A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. 63 (2), 81–97. https://doi.org/10.1037/h0043158.

Norman, Donald A., 2002. The Design of Everyday Things. Basic Books, Inc, USA.

Patterson, Robert E., Blaha, Leslie M., Grinstein, Georges G., Liggett, Kristen K., Kaveney, David E., Sheldon, Kathleen C., Havig, Paul R., Moore, Jason A., 2014. A human cognition framework for information visualization. Comput. Graph. 42, 42–58. https://doi.org/10.1016/j.cag.2014.03.002.

Radkowski, Rafael, Herrema, Jordan, Oliver, James, 2015. Augmented reality-based manual assembly support with visual features for different degrees of difficulty. Int. J. Hum. -Comput. Interact. 31 (5), 337–349. https://doi.org/10.1080/10447318.2014.994194.

Schnotz, Wolfgang, 2014. Integrated Model of Text and Picture Comprehension. The Cambridge Handbook of Multimedia Learning, Second Edition.

Sigrist, Roland, Rauter, Georg, Riener, Robert, Wolf, Peter, 2013. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. Psychon. Bull. Rev. 20 (1), 21–53.

Stork, Sonja, Schubö, Anna, 2010. Human cognition in manual assembly: theories and applications. Adv. Eng. Inform. 24 (3), 320–328. https://doi.org/10.1016/j.aei.2010.05.010.

Sweller, John, 1994. Cognitive load theory, learning difficulty, and instructional design. Learn. Instr. 4 (4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5.

Sweller, John, Chandler, Paul, 1994. Why some material is difficult to learn. Cogn. Instr. 12 (3), 185–233. https://doi.org/10.1207/s1532690xci1203_1.

Sweller, John, van Merriënboer, Jeroen J.G., Paas, Fred, 2019. Cognitive architecture and instructional design: 20 years later. Educ. Psychol. Rev. 31 (2), 261–292.

Tversky, Barbara, Zacks, Jeff, Lee, Paul, Heiser, Julie, 2000. Lines, blobs, crosses and arrows: diagrammatic communication with schematic figures. Lect. Notes Comput. Sci. 1889, 221–230. https://doi.org/10.1007/3-540-44590-0_21.

Van Genuchten, Erlijn, Scheiter, Katharina, Schüler, Anne, 2012. Examining learning from text and pictures for different task types: does the multimedia effect differ for conceptual, causal, and procedural tasks? Comput. Hum. Behav. 28 (6), 2209–2218. https://doi.org/10.1016/j.chb.2012.06.028.

Van Holland, Winfried, Bronsvoort, Willem F., 2000. Assembly features in modeling and planning. Robot. Comput. -Integr. Manuf. 16 (4), 277–294. https://doi.org/10.1016/S0736-5845(00)00014-4.

Villani, Valeria, Pini, Fabio, Leali, Francesco, Secchi, Cristian, 2018. Survey on human–robot collaboration in industrial settings: safety, intuitive interfaces and applications. Mechatronics 55, 248–266. https://doi.org/10.1016/j.mechatronics.2018.02.009.

Wang, Zhuo, Bai, Xiaoliang, Zhang, Shusheng, Wang, Yang, Han, Shu, Zhang, Xiangyu, Yan, Yuxiang, Xiong, Zhishuo, 2021. User-oriented AR assembly guideline: a new classification method of assembly instruction for user cognition. Int. J. Adv. Manuf. Technol. 112 (1–2), 41–59. https://doi.org/10.1007/s00170-020-06291-w.

Wickens, Christopher D., Helton, William S., Hollands, Justin G., Banbury, Simon, 2021. Engineering Psychology and Human Performance.

Wolfe, Jeremy M., 1994. Guided search 2.0 a revised model of visual search. Psychon. Bull. Rev. 1 (2), 202–238. https://doi.org/10.3758/BF03200774.

Zahariev, Mihaela A., MacKenzie, Christine L., 2007. Grasping at 'Thin Air': multimodal contact cues for reaching and grasping. Exp. Brain Res. 180 (1), 69–84. https://doi.org/10.1007/s00221-006-0845-4.