Integrating Semantic Communication and Human Decision-Making into an End-to-End Sensing-Decision Framework

Edgar Beck, *Graduate Student, IEEE,* Hsuan-Yu Lin, Patrick Rückert, Yongping Bao, Bettina von Helversen, Sebastian Fehrler, Kirsten Tracht, Armin Dekorsy, *Senior, IEEE*

***Abstract*—As early as 1949, Weaver defined communication in a very broad sense to include all procedures by which one mind or technical system can influence another, thus establishing the idea of semantic communication. With the recent success of machine learning in expert assistance systems where sensed information is wirelessly provided to a human to assist task execution, the need to design effective and efficient communications has become increasingly apparent. In particular, semantic communication aims to convey the meaning behind the sensed information relevant for Human Decision-Making (HDM). Regarding the interplay between semantic communication and HDM, many questions remain, such as how to model the entire end-to-end sensing-decision-making process, how to design semantic communication for the HDM and which information should be provided to the HDM. To address these questions, we propose to integrate semantic communication and HDM into one probabilistic end-to-end sensing-decision framework that bridges communications and psychology. In our interdisciplinary framework, we model the human through a HDM process, allowing us to explore how feature extraction from semantic communication can best support HDM both in theory and in simulations. In this sense, our study reveals the fundamental design trade-off between maximizing the relevant semantic information and matching the cognitive capabilities of the HDM model. Our initial analysis shows how semantic communication can balance the level of detail with human cognitive capabilities while demanding less bandwidth, power, and latency.**

***Index Terms*—6G, assistance systems, human decision-making, human-machine interface, information maximization (InfoMax), machine learning, psychology, semantic communication, task-oriented communication, wireless communications**

I. Introduction

With recent breakthroughs in Machine Learning (ML), such as generative Artificial Intelligence (AI) or Natural Language Processing (NLP), assistance systems are now finding their way into everyday life [1]. For example, doctors are supported by expert assistance systems that outperform human expertise in evaluating medical image data for disease diagnosis [2]. Many assistance systems acquire information about physical, chemical, and biological processes through sensors or sensor networks and transmit it to humans for decision-making when performing specific tasks. Applications that exploit such assistance systems include remote operation concepts for production, rescue scenarios, healthcare, autonomous driving, underwater repairs, remote sensing for earth observation and swarm exploration [3]. For example, mobile robotic systems equipped with sensors can assist Human Decision-Making (HDM). All of this relies heavily on efficient and effective wireless communications, which is therefore an integral part of the entire end-to-end sensing-decision-making process.

At this point, semantic communication comes into play as it deals with the question of how information from the assistance system can be communicated more effectively to the human to improve HDM in task execution while demanding less bandwidth, power, and latency. Several research questions can be identified from this interplay:

a) How to model the end-to-end sensing-decision-making process that bridges the disciplines communications and psychology?
b) Is semantic communication suitable for providing the information needed in terms of relevance and accuracy to facilitate effective HDM? Given a task, which and how much information should semantic communication provide, i.e., how to design semantic communication for accurate HDM?
c) Given the provided semantic information, how does the HDM process impact the end-to-end sensing-decision-making process?

To address these questions, we propose integrating semantic communication and HDM into a unified probabilistic end-to-end sensing-decision framework, thereby composing all three levels described by Weaver [4]. To showcase our framework's applicability and highlight its key mechanisms, we examine a case study grounded in an empirical categorization example. As a starting point of our study, we will first reflect upon the state of the art in semantic communication and human decision-making.

This work was partly funded by the Federal State of Bremen and the University of Bremen as part of the Humans on Mars Initiative, by the German Ministry of Education and Research (BMBF) under grant 16KISK016 (Open6GHub), and by the German Research Foundation (DFG) under grant 500260669 (SCIL).

Edgar Beck and Armin Dekorsy are with the Department of Communications Engineering, University of Bremen, 28359 Bremen, Germany (e-mail: {beck, dekorsy}@ant.uni-bremen.de). Patrick Rückert and Kirsten Tracht are with the Bremen Institute for Mechanical Engineering, University of Bremen (e-mail: {rueckert, tracht}@uni-bremen.de). Hsuan-Yu Lin and Bettina von Helversen are with the Department of Psychology, Faculty of Human and Health Sciences, University of Bremen (e-mail: {hslin, b.helversen}@uni-bremen.de). Yongping Bao and Sebastian Fehrler are with the SOCIUM Research Center on Inequality and Social Policy, University of Bremen (e-mail: {yongping.bao, sebastian.fehrler}@uni-bremen.de).

A. Semantic Communication

In the 1949 review of Shannon's general theory of communication [4], Weaver introduces the idea of semantic communication with regard to both humans and technical systems. There, he used the term communication *"in a very broad sense to include all of the procedures by which one mind may affect another. This, of course, involves not only written and oral speech, but also music, the pictorial arts, the theatre, the ballet, and in fact all human behavior. In some connections it may be desirable to use a still broader definition of communication, namely, one which would include the procedures by means of which one mechanism [...] affects another mechanism [...]."* To meet the unprecedented demands of 6G communication efficiency in terms of bandwidth, latency, and power, attention has been drawn to the broad concept of semantic communication [4]–[9]. It aims to transmit the meaning of a message rather than its exact version, which has been the focus of digital error-free system design [4]. Approaches to the description or design of semantic communication can be divided into statistical probability-based [10], logical probability-based [11], knowledge graph-based [12], and kernel-based [13].

Arguing for the generality of Shannon's theory not only for the technical level but for the semantic level design as Weaver [4], Bao, Basu et al. [14], [15] were the first to define semantic information sources and channels to tackle the semantic design by information-theoretic approaches.

With the rise of Machine Learning (ML) in communication research, transformer-based Deep Neural Networks (DNNs) have been introduced to AutoEncoders (AEs) for text transmission to learn compressed hidden representations of semantic content, aiming to improve communication efficiency [16]. However, accurate recovery of the source (text) is the main goal. The approach improves performance in semantic metrics, especially at low Signal-to-Noise Ratio (SNR), compared to classical digital transmissions. It has been adapted to many other problems, e.g., speech transmission [17], [18]. Meanwhile, also recent advances in large AI models have found their way into semantic communication [19], [20].

From a theoretical perspective, building upon the ideas of Bao, Basu et al. [14], [15], in [9], [21], the authors explicitly define a semantic random variable and identify the Information Maximization (InfoMax) problem and its variation, the Information Bottleneck (IB) problem, as appropriate semantic design criteria. Solving the InfoMax problem with ML tools, the authors obtain their design Semantic INFOrmation TraNsmission and RecoverY (SINFONY). For more details, we refer the reader to Sec. II-B and Sec. III-C. Furthermore, task-oriented edge-cloud transmission has been formulated as an IB problem [10].

Semantic communication has been extended to process several types of data, i.e., multimodal data, such as image, text, depth map data [22], [23]. In addition, monitoring, planning, and control of real worlds require the processing of multiple tasks. Thus, in [24], [25], the authors extend the concept of a semantic source to include multiple semantic interpretations. To facilitate cooperative multitask processing and improve training convergence, the semantic encoders are divided into common and specific units, extracting common low-level features and separate high-level features.

So far, the human behind the application or task has only been taken into account by theory, with the rate-distortion-perception trade-off [26], [27]. For example, the mean square error distortion is known to be inconsistent with human perception and thus not a good semantic optimization criterion [26]. Precisely because humans make the final decision when performing a task, we aim to fill the research gap of bridging semantic communication and human decision-making into an end-to-end sensing-decision framework.

B. Human Decision-Making

Even though the decision capability of artificial systems is increasing, in many situations the final decision-maker will be a human, and humans do not always make rational decisions. Therefore, to optimize the results, the needs, and capabilities of the decision-maker must be considered in the semantic communication design, e.g., by definition of the semantic source.

Humans are undoubtedly expert decision-makers who can cope well with uncertainty and complexity [28], [29]. However, it has been repeatedly shown that Human Decision-Making (HDM) can be systematically biased and decisions can be influenced by irrelevant information and context, as shown in the large literature on heuristics and biases [30]. For example, judges' sentencing decisions can be systematically influenced by asking whether a sentence should be higher or lower than a randomly generated number [31], and decisions differ depending on whether the same information is presented in frequencies or percentages [32].

Rational models of decision-making typically require the decision-maker to consider all relevant information about the decision options and the context [33]. However, humans have limited cognitive resources, such as attention or working memory capacity, which restricts the amount of information they can process at once [34], [35].

It is often assumed that humans deal with these limited capacities by using simplified decision strategies that often consider only a subset of the information and discard "extra" information [30], [33], [35]. For example, the "take-the-best" heuristic assumes that the decision-maker considers only one dimension at a time in the order of validity of the dimension. A decision is made when the decision-maker encounters a dimension that discriminates between alternatives [36]. Importantly, the use of heuristics such as the take-the-best heuristic often leads to decision performance on par with complex decision rules if the most valid predictors are indeed considered [37].

However, humans are not always able to identify the best predictors, especially when the information environment is complex, they lack expertise, are pressed for time, or are distracted [38]. In these situations, as the growing literature on decision-support/assistance systems shows, human decision-making can be supported and improved by highlighting relevant information, providing summary information, or reducing irrelevant information [39]. Even when the human decision-maker has access to all relevant information and is able to integrate the information properly, humans have a tendency to respond probabilistically [40], [41]. When given several options, and each option has a certain probability of being correct, the optimal decision (that has the highest chance of being correct) is to deterministically choose the option with the highest probability of being correct. While humans choose the best option in the majority of the trials, they usually also tend to choose other options. This variability in human decision-making has likely multiple causes [42].

Semantic communication offers the flexibility to adapt the transmitted information to facilitate the achievement of the human decision-maker's goals. The integration of semantic communication and human decision-making leads to a paradigm shift that includes the communication chain in the decision-support/assistance system.

C. Main Contributions

The main contributions addressing the above-mentioned research questions are the following:

- In this article, we propose a probabilistic end-to-end sensing-decision framework that wirelessly links sensed data with relevant information-based Human-Decision Making (HDM) by semantic communication.
- Based on this framework, we extend the information-theoretic view on semantic communication towards presentation design and HDM model training, revealing the fundamental presentation or semantic communication design trade-off between maximizing the relevant semantic information and matching the cognitive capabilities of the HDM model. In this sense, our study provides new insights for the design/interaction of semantic communication with models of HDM.
- To showcase our framework's applicability and investigate its key mechanisms, we examine a categorization example using effective HDM models. Simulation results show that, when balancing the design trade-off between feature extraction in semantic communication and cognitive constraints of the HDM model, adjusting the level of detail to match human cognitive capabilities is more important for achieving high decision accuracy than simply providing more relevant information. Moreover, uncertainty in the HDM process decreases accuracy.
- Semantic communication is able to provide the HDM model with sufficient information for making accurate decisions, while demanding less bandwidth, power, and latency compared to classical digital Shannon-based approaches.
- Finally, we provide an outlook on open research questions of our approach, including the design of information presentation through visualization, as well as game theory perspectives on sender-receiver conflicts of interest.

II. End-to-End Sensing-Decision Framework

To elaborate on our idea, we now describe our proposed end-to-end sensing-decision framework, which consists of multiple steps, exemplarily sketched in Fig. 1 and modeled as shown in Fig. 2. It is based on the semantic communication

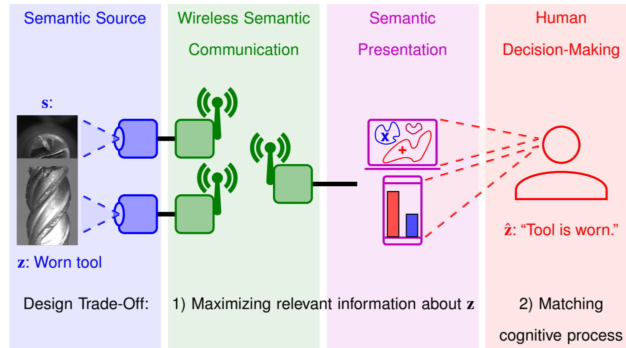

Fig. 1. Sketch of the end-to-end sensing-decision process for the example of tool wear assessment. It also situates the fundamental design trade-off between semantic communication and human decision-making.

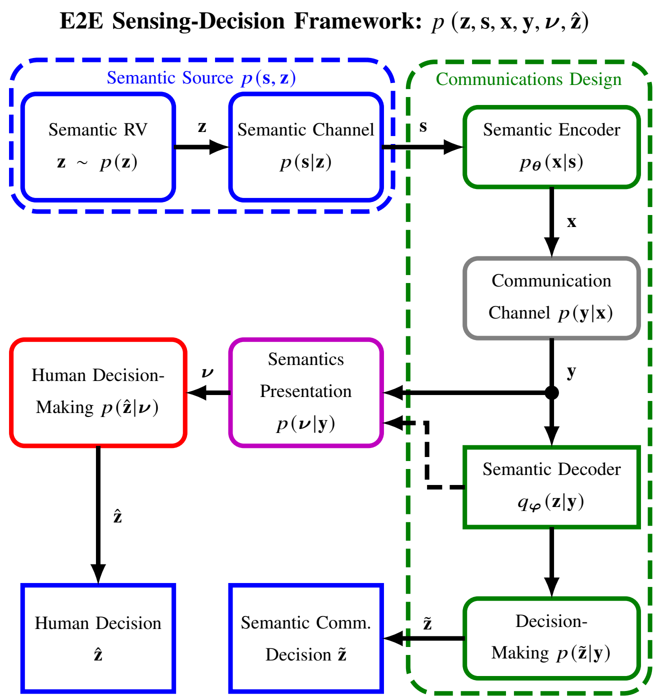

Fig. 2. Block diagram of the end-to-end sensing-decision framework, i.e., the probabilistic system model including human decision-making.

model of [9], including the complete communication Markov chain with the HDM model.

A. Semantic Source

The human performs tasks such as ensuring that machines in production run smoothly, which requires judging whether a tool is damaged or still operational. We will refer to this tool categorization task as our guiding example, whose flow is sketched in Fig. 1. The task defines the model of the world, i.e., the semantics, and is described by a semantic multivariate Random Variable (RV) $\mathbf{z} \in \mathcal{M}_z^{N_z \times 1}$ from the domain $\mathcal{M}_z$ of dimension N_z, distributed according to a probability density or mass function (pdf/pmf) $p(\mathbf{z})$ [9]. To simplify the discussion, we assume that it is discrete and memoryless.[1] The semantic source $p(\mathbf{s}, \mathbf{z})$ links the semantics expressed by $\mathbf{z}$ with the sensed, observed signal RV $\mathbf{s} \in \mathcal{M}_s^{N_s \times 1}$ that enters the communication system. This sensed data can be, e.g., images of a tool taken from different perspectives, as shown in Fig. 1. The semantic link can be modeled in the Markov chain by a semantic channel, a conditional distribution $p(\mathbf{s}|\mathbf{z})$, as shown in Fig. 2.

B. Semantic Communication

The semantic communication system encodes the sensed signal $\mathbf{s}$ with the encoder $p_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{s})$, parametrized by $\boldsymbol{\theta} \in \mathbb{R}^{N_\theta \times 1}$, to the transmit signal $\mathbf{x} \in \mathcal{M}_x^{N_{\mathrm{Tx}} \times 1}$ (see Fig. 2) for efficient and reliable semantic transmission over the physical communication channel $p(\mathbf{y}|\mathbf{x})$, where $\mathbf{y} \in \mathcal{M}_y^{N_{\mathrm{Rx}} \times 1}$ is the received signal vector, so that the semantic RV $\mathbf{z}$ is best preserved [9]. At the receiver side, the decoder $q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})$ with parameters $\boldsymbol{\varphi} \in \mathbb{R}^{N_\varphi \times 1}$ recovers the semantics $\mathbf{z}$ for the receiver.

In [9], the authors identified the Information Maximization (InfoMax) problem as an appropriate design criterion for semantic communication, since it maximizes the amount of mutual information $I_{\boldsymbol{\theta}}(\mathbf{z}; \mathbf{y})$ of the semantic RV $\mathbf{z}$ contained in the received signal $\mathbf{y}$:

$$
\begin{aligned}
& \underset{p_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{s})}{\arg\max}\ I_{\boldsymbol{\theta}}(\mathbf{z}; \mathbf{y}) && (1)\\
= \ & \underset{\boldsymbol{\theta}}{\arg\max}\ \mathrm{E}_{\mathbf{z},\mathbf{y} \sim p_{\boldsymbol{\theta}}(\mathbf{z},\mathbf{y})}\left[\ln \frac{p_{\boldsymbol{\theta}}(\mathbf{z}, \mathbf{y})}{p(\mathbf{z}) p_{\boldsymbol{\theta}}(\mathbf{y})}\right] && (2)\\
= \ & \underset{\boldsymbol{\theta}}{\arg\max}\ \mathcal{H}(\mathbf{z}) - \mathcal{H}_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{y}) && (3)\\
= \ & \underset{\boldsymbol{\theta}}{\arg\max}\ \mathrm{E}_{\mathbf{z},\mathbf{y} \sim p_{\boldsymbol{\theta}}(\mathbf{z},\mathbf{y})}\left[\ln p_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{y})\right] . && (4)
\end{aligned}
$$

There, $\mathrm{E}_{\mathbf{x} \sim p(\mathbf{x})}[f(\mathbf{x})]$ denotes the expected value of $f(\mathbf{x})$ with respect to both discrete and continuous RV $\mathbf{x}$, $\mathcal{H}(\mathbf{z}) = \mathrm{E}_{\mathbf{z} \sim p(\mathbf{z})}[-\ln p(\mathbf{z})]$ the entropy of $\mathbf{z}$, and $\mathcal{H}(\mathbf{z}|\mathbf{y})$ the conditional entropy.

If the computation of the posterior $p_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{y})$ in (4) is intractable, we can replace it by a variational distribution, i.e., the decoder $q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})$, to define a Mutual Information Lower Bound (MILBO) [9], [44], [45]:

$$
\begin{aligned}
I_{\boldsymbol{\theta}}(\mathbf{z}; \mathbf{y}) &\geq \mathcal{H}(\mathbf{z}) + \mathrm{E}_{\mathbf{z},\mathbf{y} \sim p_{\boldsymbol{\theta}}(\mathbf{z},\mathbf{y})}\left[\ln q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})\right] && (5)\\
&= \mathcal{H}(\mathbf{z}) + \mathrm{E}_{\mathbf{y} \sim p_{\boldsymbol{\theta}}(\mathbf{y})}\left[\mathrm{E}_{\mathbf{z} \sim p_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{y})}\left[\ln q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})\right]\right] && (6)\\
&= \mathcal{H}(\mathbf{z}) - \mathcal{L}_{\boldsymbol{\theta},\boldsymbol{\varphi}}^{\mathrm{CE}} . && (7)
\end{aligned}
$$

Noting that only the negative amortized cross-entropy $\mathcal{L}_{\boldsymbol{\theta},\boldsymbol{\varphi}}^{\mathrm{CE}}$ in (7) depends on both $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$ and fixing the transmit dimension to N_{Tx}, we can optimize encoder and decoder parameters [9]:

$$
\{\boldsymbol{\theta}^*, \boldsymbol{\varphi}^*\} = \underset{\boldsymbol{\theta},\boldsymbol{\varphi}}{\arg\min}\ \mathcal{L}_{\boldsymbol{\theta},\boldsymbol{\varphi}}^{\mathrm{CE}} . \tag{8}
$$

[1]For the rest of the article, note that the domain of all RVs $\mathcal{M}$ can be either discrete or continuous. Also note that the definition of entropy is different for discrete and continuous RVs. For example, the differential entropy of continuous RVs can be negative, while the entropy of discrete RVs is always positive [43]. Thus, without loss of generality, we will assume that all RVs are either discrete or continuous. In this paper, we avoid notational clutter by using the expectation operator: By replacing the integral with summation over discrete RVs, the equations are valid for continuous RVs and vice versa [9].

Note that the form of $p_{\boldsymbol{\theta}}(\mathbf{y}|\mathbf{s})$ must be constrained to avoid learning a trivial identity mapping $\mathbf{y} = \mathbf{s}$. In fact, we constrain the optimization and information rate by our communication channel $p(\mathbf{y}|\mathbf{x})$ and number of channel uses N_{Tx}, which we assume to be given. This introduces an Information Bottleneck (IB). Alternatively, we can explicitly constrain the information rate $I_{\boldsymbol{\theta}}(\mathbf{s}; \mathbf{y})$ in an IB problem [9], [21]. To solve (8), we use the empirical cross-entropy and ML techniques such as DNNs, stochastic gradient descent, and the reparametrization trick to obtain our ML-based design Semantic INFOrmation TraNsmission and RecoverY (SINFONY) [9], [21].

We note that the semantic communication system is able to make a decision by itself after optimization/training based on the decoder $q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})$. For the discrete RVs, the most likely option, i.e., the Maximum A-Posteriori (MAP) estimate, is optimal:

$$
\tilde{\mathbf{z}} = \underset{\mathbf{z}}{\arg\max}\ q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y}) . \tag{9}
$$

This decision process in operation mode is modeled as $p(\tilde{\mathbf{z}}|\mathbf{y})$.

C. Semantics Presentation

Finally, semantic communication presents the received signal $\mathbf{y}$ or the extracted probabilistic semantic decoder estimate $q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})$ — in the best case containing maximum amount of information about the semantic RV $\mathbf{z}$ according to (4) or (5) — to the HDM model. We describe this process by $p(\boldsymbol{\nu}|\mathbf{y})$ with a presentation RV $\boldsymbol{\nu} \in \mathbb{R}^{N_\phi \times 1}$. In practice, the presentation $\boldsymbol{\nu}$ must be tailored to a human, requiring a Human-Machine Interface (HMI) that is typically designed handcrafted, such as visualization (see Fig. 1). In this work, we abstract the HMI as in a technical system — where the components are connected by a deterministic function $\boldsymbol{\nu} = f(\mathbf{y})$.

D. Human Decision-Making Model

Based on the HMI or semantics presentation $\boldsymbol{\nu}$, the human decision-maker then makes a decision to complete the overall task. In this work, we will model the Human Decision-Making (HDM) process probabilistically by $p(\hat{\mathbf{z}}|\boldsymbol{\nu})$ to make a first step towards integrating and evaluating the human with the technical system, i.e., semantic communication and HDM. Finally, by decision, we obtain the estimated semantics $\hat{\mathbf{z}} \in \mathcal{M}_z^{N_z \times 1}$, which can be different from the true semantic RV $\mathbf{z}$ (see Fig. 1). In our guiding example, this could mean that the HDM process decides that the tool is damaged even though it is still usable, and vice versa.

Reflecting the variance in decision tasks, the literature on HDM includes a variety of theoretical models and approaches to capture decision processes [46], [47]. The most appropriate model often varies depending on the type of decision task and context. In this example, we focus on categorization tasks where the decision-maker must decide based on their experience whether an object belongs to one of M categories, such as whether a tool can still be used or whether the concentration of a toxic gas is above a certain threshold.

1) Generalized Context Model: While a large number of increasingly complex models of human categorization have been proposed [48]–[50], the core assumptions of the Generalized Context Model (GCM) [51], [52] are commonly adapted by many successors [53] and have been successfully used to describe categorizations of complex real world stimuli [54], [55].

Despite the relative simplicity of the GCM, the well-studied model and its variants can easily account for HDM under different contexts, e.g., under time pressure [56], [57], capture human judgment under cognitive load [58], and explain common HDM biases, e.g., base-rate bias [59]. At the same time, GCM has been applied to different aspects of human cognition, e.g., artificial grammar [60], leadership competence judgment [61], and mental multiplication [62]. Several extensions of GCM have been created to explain even broader aspects of HDM such as reaction time in decision-making [63] or learning [64]. Since the GCM is a powerful approximation to human categorization decisions, we choose it to simulate the decision-making process.

An important assumption of the GCM is that categorization decisions are made on the basis of exemplar memory, i.e., previously seen realizations that are retrieved from memory. Accordingly, in the tool example, the model assumes that the decision-maker first experiences N tool realizations and whether those tools need to be replaced. These tools are then remembered as the i-th "exemplar", i.e., realization $\boldsymbol{\nu}_i$, with the corresponding label z_i, so we have an exemplar dataset or HDM knowledge base $\mathcal{D}_{\mathrm{HK}} = \{\boldsymbol{\nu}_i, \mathsf{z}_i\}_{i=1}^{N}$. Since the semantic RV $\mathbf{z}$ is a categorical RV, we can describe it by a one-hot vector $\mathbf{z} = \text{one-hot}(k)$ where all elements are zero except for the element $k \in \{1, \ldots, M\}$ that represents the tool state from a total number of M states. For example, for binary states, we have $k \in \{1, 2\}$ with $M = 2$.

When the decision-maker encounters a new tool presentation $\boldsymbol{\nu}$, the probability $q_{\varphi_{\mathrm{G}}}(\mathbf{z} = \mathsf{z}|\boldsymbol{\nu}, \mathcal{D}_{\mathrm{HK}}) = q(\mathbf{z} = \mathsf{z}|\boldsymbol{\nu}, \mathcal{D}_{\mathrm{HK}}, \varphi_{\mathrm{G}})$ of making the decision $\mathsf{z} = \text{one-hot}(\mathsf{k})$ given this representation $\boldsymbol{\nu}$ is the result of the comparison between $\boldsymbol{\nu}$ and all seen realizations $\boldsymbol{\nu}_i$ from $\mathcal{D}_{\mathrm{HK}}$:

$$q_{\varphi_{\mathrm{G}}}(\mathbf{z} = \mathsf{z}|\boldsymbol{\nu}, \mathcal{D}_{\mathrm{HK}}) = \frac{\sum_{i=1}^{N} \mathrm{sim}(\boldsymbol{\nu}_i, \boldsymbol{\nu}|\varphi_{\mathrm{G}}) \cdot [\mathsf{z}_i = \mathsf{z}]}{\sum_{i=1}^{N} \mathrm{sim}(\boldsymbol{\nu}_i, \boldsymbol{\nu}|\varphi_{\mathrm{G}})} \tag{10}$$

with GCM parameters φ_{G} and $[\mathsf{z}_i = \mathsf{z}]$ being the Iverson bracket, which is equal to 1 if $\mathsf{z}_i = \mathsf{z}$ and 0 otherwise. This means the approximating posterior $q_{\varphi_{\mathrm{G}}}(\mathbf{z}|\boldsymbol{\nu}, \mathcal{D}_{\mathrm{HK}})$ is determined by the sum of similarities between $\boldsymbol{\nu}$ and all the seen realizations z_i that belong to the decision z, and normalized by the similarity to all the seen realizations regardless of the decision. We note that the model (10) assumes that the decision-maker has perfect memory of its knowledge base $\mathcal{D}_{\mathrm{HK}}$.

The similarity $\mathrm{sim}(\boldsymbol{\nu}_1, \boldsymbol{\nu}_2|\varphi_{\mathrm{G}})$ between two presentations decreases exponentially as the Euclidean distance between two presentations increases and depends on the learnable GCM parameters $\varphi_{\mathrm{G}} = \{\gamma, \mathbf{w}\}$:

$$\mathrm{sim}(\boldsymbol{\nu}_1, \boldsymbol{\nu}_2|\varphi_{\mathrm{G}}) = e^{-\gamma \cdot \left(|\boldsymbol{\nu}_1 - \boldsymbol{\nu}_2|^T \cdot \mathrm{diag}\{\mathbf{w}\} \cdot |\boldsymbol{\nu}_1 - \boldsymbol{\nu}_2|\right)^{1/2}} \tag{11}$$

with $\mathrm{diag}\{\mathbf{w}\}$ creating a diagonal matrix with elements of $\mathbf{w} \in \mathbb{R}^{N_\phi \times 1}$ on its diagonal.

A unique assumption of the GCM is that each element or feature ν_n of a seen realization $\boldsymbol{\nu}_i$ is weighted by attention weights w_n and hence does not contribute equally to the perceived similarity. To achieve the attention functionality, the weights are constrained by $w_n \geq 0$ and normalized by $\sum_{n=1}^{N_\phi} w_n = 1$. The parameter γ has two interpretations: First, the similarity gradient γ describes the sharpness of the decline in similarity, with higher γ resulting in a sharper decline of similarity when the distance increases. Second, the parameter γ describes the consistency in making decisions and reflects the probabilistic nature of the HDM process, akin to the temperature parameter in the Boltzmann distribution [65]. Accordingly, a lower parameter γ results in a new tool being more confidently categorized to the category with higher similarity.

2) HDM-based Probabilistic Decision-Making: After training of the GCM (see Sec. II-F.1), the strategy of the HDM model is equal to the random process

$$\hat{\mathbf{z}} \sim p(\hat{\mathbf{z}}|\boldsymbol{\nu}) = q_{\varphi_{\mathrm{G}}}(\mathbf{z} = \hat{\mathbf{z}}|\boldsymbol{\nu}, \mathcal{D}_{\mathrm{HK}}) \, . \tag{12}$$

When faced with options of varying probabilities, people tend to distribute their choices according to the probability distribution rather than always selecting the most likely option resulting in a suboptimal policy. Thus, the accuracy of human decisions can be worse compared to that of the optimal deterministic policy (9) of the technical (semantic communication) system, e.g., SINFONY.

E. End-to-End Sensing-Decision Model

With all the aforementioned subcomponent models, we are able to create a generative model of the end-to-end sensing-decision process. We note that we can distinguish between four different models corresponding to four system stages:

1) Design of semantic encoder $p_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{s})$ and decoder $q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})$ based on the forward communications model
$$p(\mathbf{z}, \mathbf{s}, \mathbf{x}, \mathbf{y}) = p(\mathbf{z}, \mathbf{s}) \cdot p_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{s}) \cdot p(\mathbf{y}|\mathbf{x}) \, . \tag{13}$$
2) Semantic communication is executed in operation mode to make decisions via (9). Then, the model is
$$p(\mathbf{z}, \mathbf{s}, \mathbf{x}, \mathbf{y}, \tilde{\mathbf{z}}) = p(\mathbf{z}, \mathbf{s}, \mathbf{x}, \mathbf{y}) \cdot p(\tilde{\mathbf{z}}|\mathbf{y}) \, . \tag{14}$$
3) The HDM model $q_{\varphi_{\mathrm{G}}}(\mathbf{z}|\boldsymbol{\nu}, \mathcal{D}_{\mathrm{HK}})$ is trained based on seen presentation and label realizations from semantic communication in operation mode. The underlying model is
$$p(\mathbf{z}, \mathbf{s}, \mathbf{x}, \mathbf{y}, \boldsymbol{\nu}) = p(\mathbf{z}, \mathbf{s}, \mathbf{x}, \mathbf{y}) \cdot p(\boldsymbol{\nu}|\mathbf{y}) \, . \tag{15}$$
4) Semantic communication presents information to the HDM model that finally makes a decision. The overall end-to-end sensing-decision model of Fig. 2 in operation mode after all training phases is:
$$p(\mathbf{z}, \mathbf{s}, \mathbf{x}, \mathbf{y}, \boldsymbol{\nu}, \hat{\mathbf{z}}) = p(\mathbf{z}, \mathbf{s}, \mathbf{x}, \mathbf{y}, \boldsymbol{\nu}) \cdot p(\hat{\mathbf{z}}|\boldsymbol{\nu}) \, . \tag{16}$$

F. Information-theoretic Overall View on Design in the End-to-End Sensing-Decision Framework

We can exploit our end-to-end sensing-decision framework (13)-(16), to extend the information-theoretic view of semantic communication to both the semantics presentation and the HDM model to gain new insights.

1) HDM Model – Training: For optimization of the GCM parameters $\boldsymbol{\varphi}_{\mathrm{G}} = \{\gamma, \mathbf{w}\}$ given a presentation $\boldsymbol{\nu}$ based on a fixed optimized semantic communication system of model (15), typically the maximum likelihood criterion is used [52]. We note that maximization of the log-likelihood function is equal to amortized minimization of the empirical cross-entropy on the training set [43]. Transferring the InfoMax view from the semantic communication system in (7), this means we optimize a lower bound on the mutual information $I_{\boldsymbol{\theta}}(\mathbf{z};\boldsymbol{\nu})$, but now between $\mathbf{z}$ and $\boldsymbol{\nu}$ with respect to $\boldsymbol{\varphi}_{\mathrm{G}}$:

$$
\begin{aligned}
I_{\boldsymbol{\theta}}(\mathbf{z};\boldsymbol{\nu}) &\geq \mathcal{H}(\mathbf{z}) + \mathrm{E}_{\mathbf{z},\boldsymbol{\nu}\sim p_{\boldsymbol{\theta}}(\mathbf{z},\boldsymbol{\nu})}\left[\ln q_{\boldsymbol{\varphi}_{\mathrm{G}}}(\mathbf{z}|\boldsymbol{\nu},\mathcal{D}_{\mathrm{HK}})\right] \quad (17)\\
&= \mathcal{H}(\mathbf{z}) - \mathcal{L}^{\mathrm{CE}}_{\boldsymbol{\varphi}_{\mathrm{G}}}\,. \quad (18)
\end{aligned}
$$

From an information-theoretic view, we conclude that the choice of the GCM optimization criterion

$$
\boldsymbol{\varphi}^{*}_{\mathrm{G}} = \underset{\boldsymbol{\varphi}_{\mathrm{G}}}{\arg\min}\ \mathcal{L}^{\mathrm{CE}}_{\boldsymbol{\varphi}_{\mathrm{G}}} \quad (19)
$$

is well-motivated. To solve (19), computer search methods are typically used [52]. In this work, we employ a variant known as the differential annealing algorithm. By integrating differential evolution's population-based search with simulated annealing's probabilistic acceptance of solutions, differential annealing aims to balance exploration and exploitation in complex search spaces to enhance global optimization capabilities [66].

2) Semantics Presentation – Design Optimization: Moreover, if we add the presentation process as a tunable encoder $p_{\boldsymbol{\theta}_{\mathrm{P}}}(\boldsymbol{\nu}|\mathbf{y})$ with parameters $\boldsymbol{\theta}_{\mathrm{P}}$ to the optimization problem (19), we arrive at the MILBO objective function:

$$
\begin{aligned}
I_{\boldsymbol{\theta},\boldsymbol{\theta}_{\mathrm{P}}}(\mathbf{z};\boldsymbol{\nu}) \geq& \mathcal{H}(\mathbf{z})\\
&+ \mathrm{E}_{\mathbf{z},\mathbf{y},\boldsymbol{\nu}\sim p_{\boldsymbol{\theta}}(\mathbf{z},\mathbf{y})\cdot p_{\boldsymbol{\theta}_{\mathrm{P}}}(\boldsymbol{\nu}|\mathbf{y})}\left[\ln q_{\boldsymbol{\varphi}_{\mathrm{G}}}(\mathbf{z}|\boldsymbol{\nu},\mathcal{D}_{\mathrm{HK}})\right]\\
=& \mathcal{H}(\mathbf{z}) - \mathcal{L}^{\mathrm{CE}}_{\boldsymbol{\theta}_{\mathrm{P}},\boldsymbol{\varphi}_{\mathrm{G}}}\,. \quad (20)
\end{aligned}
$$

Decomposing the amortized cross-entropy $\mathcal{L}^{\mathrm{CE}}_{\boldsymbol{\theta}_{\mathrm{P}},\boldsymbol{\varphi}_{\mathrm{G}}}$ as in [9] into

$$
\begin{aligned}
\mathcal{L}^{\mathrm{CE}}_{\boldsymbol{\theta}_{\mathrm{P}},\boldsymbol{\varphi}_{\mathrm{G}}} &= \mathcal{H}(\mathbf{z}) - I_{\boldsymbol{\theta},\boldsymbol{\theta}_{\mathrm{P}}}(\mathbf{z};\boldsymbol{\nu}) \quad (21)\\
&+ \mathrm{E}_{\boldsymbol{\nu}\sim p_{\boldsymbol{\theta},\boldsymbol{\theta}_{\mathrm{P}}}(\boldsymbol{\nu})}\left[D_{\mathrm{KL}}\left(p_{\boldsymbol{\theta},\boldsymbol{\theta}_{\mathrm{P}}}(\mathbf{z}|\boldsymbol{\nu})\parallel q_{\boldsymbol{\varphi}_{\mathrm{G}}}(\mathbf{z}|\boldsymbol{\nu},\mathcal{D}_{\mathrm{HK}})\right)\right]
\end{aligned}
$$

reveals two possibly conflicting design criteria:

1) The presentation encoder $p_{\boldsymbol{\theta}_{\mathrm{P}}}(\boldsymbol{\nu}|\mathbf{y})$ should maximize the mutual information $I_{\boldsymbol{\theta},\boldsymbol{\theta}_{\mathrm{P}}}(\mathbf{z};\boldsymbol{\nu})$ that depends solely on it through the true posterior $p_{\boldsymbol{\theta},\boldsymbol{\theta}_{\mathrm{P}}}(\mathbf{z}|\boldsymbol{\nu})$ (see (4)).
2) Both true posterior $p_{\boldsymbol{\theta},\boldsymbol{\theta}_{\mathrm{P}}}(\mathbf{z}|\boldsymbol{\nu})$ and hence the presentation encoder $p_{\boldsymbol{\theta}_{\mathrm{P}}}(\boldsymbol{\nu}|\mathbf{y})$ and the HDM model $q_{\boldsymbol{\varphi}_{\mathrm{G}}}(\mathbf{z}|\boldsymbol{\nu},\mathcal{D}_{\mathrm{HK}})$ are matched by minimizing the Kullback-Leibler (KL) divergence.

In a technical semantic communication system from Sec. II-B, we can avoid the design conflict by using a model $q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})$ expressive enough to approximate $p_{\boldsymbol{\theta}}(\mathbf{z}|\mathbf{y})$ arbitrarily well, such that the focus lies on the InfoMax term. However, in case of the end-to-end sensing-decision training model (15), if the HDM model (or human) constrains the form of $q_{\boldsymbol{\varphi}_{\mathrm{G}}}(\mathbf{z}|\boldsymbol{\nu},\mathcal{D}_{\mathrm{HK}})$, i.e., the solution space, to some degree, the two optimization terms in (21) are traded-off: Then, the true posterior $p_{\boldsymbol{\theta},\boldsymbol{\theta}_{\mathrm{P}}}(\mathbf{z}|\boldsymbol{\nu})$ has to be fit to $q_{\boldsymbol{\varphi}_{\mathrm{G}}}(\mathbf{z}|\boldsymbol{\nu},\mathcal{D}_{\mathrm{HK}})$ and we do not maximize $I_{\boldsymbol{\theta},\boldsymbol{\theta}_{\mathrm{P}}}(\mathbf{z};\boldsymbol{\nu})$ alone which could lead to a loss in mutual information.

Example 1: These abstract information-theoretic insights explain well what we observe in practice with handcrafted presentations. In reality, it is difficult to understand and subsequently visualize the received raw communications signal $\mathbf{y}$ for a human without any preprocessing:

- We have to match the presentation encoder $p_{\boldsymbol{\theta}_{\mathrm{P}}}(\boldsymbol{\nu}|\mathbf{y})$ to the cognitive process $q_{\boldsymbol{\varphi}_{\mathrm{G}}}(\mathbf{z}|\boldsymbol{\nu},\mathcal{D}_{\mathrm{HK}})$ according to the KL divergence term in (21). Fortunately, the semantic decoder $q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})$ obtained by maximizing the MILBO extracts the semantic information of $\mathbf{z}$ from $\mathbf{y}$ and allows for meaningful presentation to and interpretation by the HDM model $q_{\boldsymbol{\varphi}_{\mathrm{G}}}(\mathbf{z}|\boldsymbol{\nu},\mathcal{D}_{\mathrm{HK}})$ (or human).
- However, we may lose relevant information about $\mathbf{z}$ fitting the presentation to the cognitive process including the semantic decoder preprocessing in the Markov chain $\mathbf{z}\rightarrow\mathbf{y}\rightarrow q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})\rightarrow\boldsymbol{\nu}$ according to the data processing inequality

$$
I(\mathbf{z};\mathbf{y}) \geq I\left(\mathbf{z};q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})\right) \geq I(\mathbf{z};\boldsymbol{\nu})\ . \quad (22)
$$

Example 2: Another example of how the HDM process influences presentation design is that research on HDM models focuses on the interplay between relevant features $\boldsymbol{\nu}$ and require certain level of feature extraction from raw images $\mathbf{s}$ (or $\mathbf{y}$) for HDM model processing. For an overview of this research, we refer the reader to [67]. These HDM models were not built to process raw images $\mathbf{s}$ directly, i.e., $q_{\boldsymbol{\varphi}_{\mathrm{G}}}(\mathbf{z}|\boldsymbol{\nu}=\mathbf{s},\mathcal{D}_{\mathrm{HK}})$, which would lead to unrealistically poor performance despite maximum relevant information in $\mathbf{s}$ about $\mathbf{z}$. Thus, in the numerical results of Sec. III-D, we cannot compare to a setup where the raw data of the images $\mathbf{s}$ are digitally communicated and then directly processed by the HDM model.

Based on our end-to-end sensing-decision framework, we conclude that it highly depends on the processing capabilities of the HDM model if it can extract more or less information about $\mathbf{z}$ from $\mathbf{y}$ than the semantic decoder. Moreover, we conclude that balancing of two possibly conflicting criteria is key for presentation design:

1) **Relevant information preservation:** On the one hand, careful design of $\boldsymbol{\nu}$ is required to not lose any relevant information about $\mathbf{z}$ for the final decision. For example, the higher the dimension N_{ϕ} of the presentation RV $\boldsymbol{\nu}$, the more detailed the presentation to the HDM model and the more information it contains.
2) **Presentation alignment to the HDM model:** On the other hand, the presentation has to be in a form that can be understood by the HDM model, effectively restricting the set of possible presentations $\boldsymbol{\nu}$. For example, compressing the relevant information about $\mathbf{z}$ into $\boldsymbol{\nu}$ may be required to ease cognitive processing.

To investigate how to balance these two design rules, we compare two handcrafted presentations in our numerical example of Sec. III-D.1. For an outlook on the inclusion of HMIs in practice, please refer to Sec. IV.

III. Simulative Investigation

In this section, we evaluate first numerical results of our joint framework using the example of image classification. The inclusion of diverse datasets for semantic source modeling, such as the standard MNIST and CIFAR10 datasets in addition to our guiding tool example, enables the generalization of conclusions beyond the specific case of tool wear.

A. Performance Measures

We measure the performance of decision-making for both semantic communication and the HDM model by the categorical accuracy

$$\mathcal{A} = \mathrm{E}_{\mathbf{z}\sim p(\mathbf{z})}\left[p\left(\tilde{\mathbf{z}} = \mathbf{z}|\mathbf{z}\right)\right] = \sum_{\mathbf{z}\in\mathcal{M}_z^{N_z\times 1}} p\left(\tilde{\mathbf{z}} = \mathbf{z}, \mathbf{z}\right) \tag{23}$$

$$\approx \frac{1}{N}\sum_{i=1}^{N}\left[\tilde{\mathbf{z}}_i = \mathbf{z}_i\right] \tag{24}$$

— comparing predicted and true category realizations $\tilde{\mathbf{z}}_i$ and $\mathbf{z}_i$ — or the classification error rate $1-\mathcal{A}$ common in communications. Since the HDM model decides probabilistically based on the input $\boldsymbol{\nu}_i$, we can calculate the accuracy based on the end-to-end sensing-decision model (16) shown in Fig. 2 by the sum of the probabilities of the GCM responding to the correct category $\mathbf{z}_i$ [51]:

$$\mathcal{A} = \mathrm{E}_{\mathbf{z}\sim p(\mathbf{z})}\left[p\left(\hat{\mathbf{z}} = \mathbf{z}|\mathbf{z}\right)\right] \tag{25}$$

$$= \mathrm{E}_{\mathbf{z}\sim p(\mathbf{z})}\left[\mathrm{E}_{\boldsymbol{\nu}\sim p(\boldsymbol{\nu}|\mathbf{z})}\left[p\left(\hat{\mathbf{z}} = \mathbf{z}|\boldsymbol{\nu}\right)\right]\right] \tag{26}$$

$$= \mathrm{E}_{\mathbf{z},\boldsymbol{\nu}\sim p(\mathbf{z},\boldsymbol{\nu})}\left[p\left(\hat{\mathbf{z}} = \mathbf{z}|\boldsymbol{\nu}\right)\right] \tag{27}$$

$$\approx \frac{1}{N}\sum_{i=1}^{N} p(\hat{\mathbf{z}} = \mathbf{z}_i|\boldsymbol{\nu} = \boldsymbol{\nu}_i) \tag{28}$$

$$= \frac{1}{N}\sum_{i=1}^{N} q_{\varphi_{\mathrm{G}}}(\mathbf{z} = \mathbf{z}_i|\boldsymbol{\nu} = \boldsymbol{\nu}_i, \mathcal{D}_{\mathrm{HK}})\,. \tag{29}$$

This method of calculating accuracy is commonly used in psychology studies for HDM models [51].

B. Example Semantic Source Datasets

Tool wear and tool replacement decisions pose a common challenge in the metal cutting industry to reduce production costs [68], and represent an exemplary semantic source of this work. In this decision-making problem, the semantic RV $\mathbf{z}$ is modeled as a binary variable with two states, where $\mathbf{z} = [1, 0]^T$ indicates a worn tool and $\mathbf{z} = [0, 1]^T$ indicates a usable tool. Although optical measurement techniques provide accurate assessments of tool wear, small and medium-sized companies often rely on machine operators to manually assess tool wear. To automate this process, a dataset was recorded where human experts were presented two different grayscale images of each tool [69]: One image $\mathbf{s}_1 \in \{0, 1, \ldots, 255\}^{218\times 380\times 1}$ taken from the side and another $\mathbf{s}_2 \in \{0, 1, \ldots, 255\}^{487\times 380\times 1}$ taken from the top (see Fig. 1). Based on these observations $\mathbf{s} = \{\mathbf{s}_1, \mathbf{s}_2\}$, the experts labeled the tools into the binary states $\mathbf{z}$.

This process resulted in a dataset $\mathcal{D} = \{\mathbf{s}_i, \mathbf{z}_i\}_{i=1}^N$ modeling our semantic source $p(\mathbf{s}, \mathbf{z})$ and consisting of $N = 1632$ data pairs, with an 85% split between training and validation data. We also revisit the MNIST and CIFAR10 examples from [9], [70] to extend our analysis.

C. Semantic Communication Analysis

As the design approach for semantic communication and to solve (8) with the model (13), we use our ML-based SINFONY approach from [9], [21]. However, we note that the conclusions derived from the results about the interplay between semantic communication and HDM models are not limited to this approach. These also extend to other, e.g., model-based, methods capable of providing the same quality of soft information at inference runtime. For example, RL-SINFONY leverages reinforcement learning to train the design via (7) to comparable performance [21].

1) SINFONY Design: As shown in [9], [21], we apply SINFONY to a distributed multipoint scenario, where meaning from multiple image sources is communicated to a single receiver for semantic recovery of the RV $\mathbf{z}$. In the numerical example of [9], four distributed agents extract features from different image views with an encoder based on the famous and powerful ResNet architecture [71] for rate-efficient transmission. Based on the received signals, the decoder recovers semantics by classification. Numerical results of [9] on images from the MNIST and CIFAR10 datasets show that SINFONY outperforms classical digital communication systems in terms of bandwidth, latency and power efficiency.

In this article, we reuse the SINFONY approach for integration with the HDM model. SINFONY is particularly well-suited for integration because it can be easily adapted to any semantic source $p(\mathbf{s}, \mathbf{z})$, i.e., use case, including tool damage classification, by changing the data samples and specifically designing its DNN architecture.

In the guiding tool example, two image sensors provide different views of the tool (see Fig. 1). This results in a SINFONY design (see Communications Design in Fig. 2) with two encoders $p_{\boldsymbol{\theta}}^i(\mathbf{x}_i|\mathbf{s}_i)$ with $i = \{1, 2\}$ that can be concatenated into one virtual encoder $p_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{s})$, and one decoder $q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})$. Owing to the large image dimensions of $\mathbf{s}_i$, we adopt the ImageNet version of ResNet18 to reduce numerical complexity in feature extraction [72], resulting in $N_{\mathrm{Feat}} = 512$ features per encoder. We test two SINFONY Tx module designs that map those features onto the transmit signal $\mathbf{x}_i \in \mathcal{M}_x^{N_{\mathrm{Tx}}\times 1}$: one with feature compression ($N_{\mathrm{Tx}} = 128$) and one without ($N_{\mathrm{Tx}} = 512$). Note that the number of channel uses N_{Tx} is proportional to bandwidth, i.e., $B \sim N_{\mathrm{Tx}}$. The signals $\mathbf{x}_i$ are transmitted over an AWGN channel $p(\mathbf{y}_i|\mathbf{x}_i)$ to the decoder that consists of a common Rx layer of width $N_{\mathrm{w}} = 1024$ processing the concatenated received signals $\mathbf{y}_i \in \mathcal{M}_y^{N_{\mathrm{Rx}}\times 1}$, each of length $N_{\mathrm{Rx}} = N_{\mathrm{Tx}}$, and a final softmax layer with $M = 2$ classes. As in [9], we train for $\mathrm{SNR}_{\mathrm{train}} \in [-4, 6]$ dB in model (13). We also reuse the SINFONY designs for the

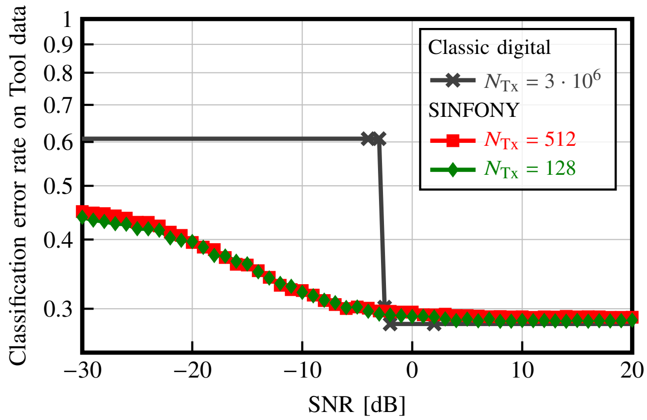

Fig. 3. Comparison of the classification error rate of SINFONY with different number of channel uses N_{Tx} per encoder and central image processing with digital image transmission on the tool validation dataset as a function of SNR.

MNIST and CIFAR10 datasets from [9] and combine them with the HDM model. The source code including all details is available in [70].

2) SINFONY-based Decision-Making: First, we evaluate the performance of SINFONY-based decision-making within the semantic communication operation mode model (14), where SINFONY makes the final decision via (9). Fig. 3 shows the classification error rate $1 - \mathcal{A}$ with $\mathcal{A}$ from (24) on the tool dataset as a function of SNR. The key findings are similar to those for MNIST and CIFAR10 [9] but become much more obvious: Using less channel uses per encoder with SINFONY $N_{\mathrm{Tx}} = 128$ than the number of features ($N_{\mathrm{Feat}} = 512$) results in the same performance compared to SINFONY Tx/Rx $N_{\mathrm{Tx}} = 512$. This indicates that feature compression and thus a reduction in bandwidth is possible.

Moreover, we compare to central image processing by a ResNet classifier [70] with classic digital transmission of the sensed images (Classic digital): We assume that the RGB image bits are Huffman encoded, protected by an LDPC code with rate 0.25 and BPSK modulated. At the receiver side, we use belief propagation for decoding. On average, the channel is utilized over 23,400 times more frequently per encoder, with $N_{\mathrm{Tx}} \approx 2,998,626.82 \approx 3 \cdot 10^6$ uses. Furthermore, at low SNR, significantly more power is needed to achieve the same classification error rate, e.g., about 10 dB more for 35%. Instead of graceful degradation as for SINFONY, we observe a cliff effect typical for digital communication at a SNR threshold of −2.5 dB: Communication quality remains robust as long as channel capacity exceeds code rate and the LDPC code operates within its working point, but rapidly breaks down otherwise. This sharp contrast in performance and bandwidth highlights the huge potential of semantic communication.

D. End-to-end Sensing-Decision Analysis

Now, we assume our end-to-end sensing-decision model, i.e., the overall model (16): The semantic information in the images is transmitted by SINFONY over an AWGN channel and then fed into the HDM model, i.e., the GCM. Note that, in contrast to the SINFONY scenario of Sec. III-C.2, the HDM model now makes the final decision.

1) Semantics Presentation Design: In Sec. II-F.2, we derive two design criteria for the semantics presentation $\boldsymbol{\nu}$: 1) It should keep all relevant information about $\mathbf{z}$. 2) It should fit to cognitive processing capabilities of the HDM model. We note that since HDM models are not capable to process the raw data of the images $\mathbf{s}$ directly [67] as outlined in Sec. II-F.2, we cannot simply compare to a setup where the raw data of the images $\mathbf{s}$ are digitally communicated and processed. This means the design choice $\boldsymbol{\nu} = \mathbf{s}$ is ruled out in this work.

Therefore, we aim to gain first insights on the design trade-off by comparing HDM model performance with different presentations that reflect a different weighting of the two design criteria. To reflect practical considerations as outlined in Sec. II-F.2, we design the presentation handcrafted based on the semantic decoder (see Fig. 1). In this context, in other words, we investigate how to balance the feature extraction of semantic communication and HDM models to achieve the best task performance, i.e., to minimize the classification error rate.

We present either the categorical probability outputs (E2E categorical) or the relevant decision features (E2E N_ϕ) from SINFONY as $\boldsymbol{\nu}$ to the HDM model, i.e., the GCM:

1) **E2E categorical:** The low-dimensional and interpretable probability estimate of SINFONY for each category (E2E categorical) that fulfills design rule 2), e.g., whether the tool is damaged or not:
$$\boldsymbol{\nu} = f_1(q_\varphi(\mathbf{z}|\mathbf{y})) = \begin{bmatrix} q_\varphi\left(\mathbf{z} = \text{one-hot}(1)^T|\mathbf{y}\right) \\ \vdots \\ q_\varphi\left(\mathbf{z} = \text{one-hot}(M)^T|\mathbf{y}\right) \end{bmatrix} . \quad (30)$$

2) **E2E $\boldsymbol{N_\phi}$:** To provide the HDM model at an abstract level with potentially more relevant semantic information about $\mathbf{z}$ according to data processing inequality (22) and design rule 1) for decision-making, we use the relevant decision features
$$\boldsymbol{\nu} = f_2(\mathbf{y}) = \mathbf{v}_{q_\varphi}^{(N_{\mathrm{L}}-1)}(\mathbf{y}) \, , \quad (31)$$
where $\mathbf{v}_{q_\varphi}^{(l)}$ is the output of the l-th layer of $q_\varphi(\mathbf{z}|\mathbf{y})$ and N_{L} the depth of the DNN. This means we extract the inputs to the final dense softmax layer of the SINFONY decoder used for probability estimation, similar to a previous study that aims to model categorization with natural material [73].
To further vary the level of detail or dimension of the presentation, we extract the most important $N_\phi = \{5, 10, 20, 40\}$ of the $N_{\mathrm{h}}^{(N_{\mathrm{L}}-1)}$ final layer features to facilitate effective processing of the HDM model according to design rule 2). The importance of the i-th feature $[\mathbf{v}_{q_\varphi}^{(N_{\mathrm{L}}-1)}]_i$, with respect to all output nodes $q_\varphi(\mathbf{z} = \text{one-hot}(k)^T|\mathbf{y})$, is quantified by the sum of the absolute weight values in each column of the last-layer weight matrix $\mathbf{W}^{(N_{\mathrm{L}}-1)}$, given by $\sum_{k=1}^{M} |w_{ki}^{(N_{\mathrm{L}}-1)}|$.

Based on the selected presentation, i.e., SINFONY features, the GCM classifies into M categories, e.g., into the binary tool states of wear and non-wear. To present the SINFONY features in the numerical evaluation, we use the SINFONY version with $N_{\mathrm{Tx}} = 128$ from Fig. 3 for semantic communication (see Sec. III-C) on the tool dataset. For MNIST and CIFAR10 datasets, we use the trained SINFONY versions with $N_{\mathrm{Tx}} = 56$ and $N_{\mathrm{Tx}} = 64$ from [9], [70]. Note that N_{Tx} differs per dataset, since we tailored the SINFONY architecture to the specific dataset.

2) Simulation Scenarios: We evaluate our proposed framework on three datasets: Tools, MNIST, and CIFAR10. Furthermore, we perform two main simulations — a simulation of the accuracy as a function of the SNR typical for communications, and a simulation of the expertise of the HDM model touching a psychological aspect:

1) In the SNR simulation, we assume that the HDM model (10) has sufficient experience with the presented SINFONY features, i.e., it has perfect memory of the training set with $\mathcal{D}_{\mathrm{HK}} = \mathcal{D}_{\mathrm{T}}$, i.e., the seen presented realizations encompass the entire training dataset of semantic communication. Its attention weights $\mathbf{w}$ and the similarity gradient γ from (11) are optimized to maximize the classification accuracy on the training set at a training SNR of 20 dB. For evaluation, the HDM model receives the output of SINFONY under varying SNR.
2) In the expert simulation, we assume the highest evaluated SNR of 20 dB during communication and vary the number of seen presentation realizations, i.e., images randomly selected from the training set. We define the number of seen realizations $|\mathcal{D}_{\mathrm{HK}}|$ of classified tools as the expertise of the HDM model and simulate the performance at this expertise independently. Accordingly, a HDM model with high expertise has a larger knowledge base $\mathcal{D}_{\mathrm{HK}}$ compared to a HDM model with low expertise. The GCM parameters were optimized for the specific HDM knowledge bases $\mathcal{D}_{\mathrm{HK}}$.

In both simulations, the accuracy was calculated based on the validation dataset. We performed multiple Monte Carlo runs for evaluation: For the SNR simulations, we iterated ten times through the dataset for each SNR value. For the expertise simulations, we iterated 100 times for each expertise level.

3) Numerical Results: We present the simulation results in Fig. 4, comparing to SINFONY-based decision-making (SINFONY) of the technical system on $\tilde{\mathbf{z}}$ via (9) as the baseline. First, we note that the accuracy of the GCM is worse than that of SINFONY. This can be explained by the probabilistic decision process (12) of the HDM model, which deviates from the optimal strategy to choose the most likely option.

SNR Simulation: Furthermore, the SNR curves show a similar trend for all three datasets. Accuracy increases as a function of SNR and plateaus after a certain SNR is reached. The performance of the HDM model is best when receiving the categorical probability input of SINFONY (E2E categorical) compared to receiving the N_ϕ most important feature dimensions (E2E N_ϕ). With the latter input, more features yield better accuracy, and the performance reaches an asymptote after a certain number of feature dimensions. The number of features needed to reach saturation varies depending on the dataset.

Reaching saturation indicates that the HDM model is not able to extract more relevant information about $\mathbf{z}$ from many feature inputs, i.e., from a more detailed representation. In contrast, the model performs better with SINFONY's preprocessed probability estimates, indicating SINFONY's ability to efficiently extract the semantic information. This indicates that InfoMax optimized outputs are suitable as an input for human decision-making: SINFONY's graceful degradation translates directly into the GCM curve.

Expertise Simulation: The expertise simulation (bottom row in Fig. 4) shows that accuracy increases with the number of seen presentation realizations. Regardless of expertise, using the probability output of SINFONY again results in the best performance compared to receiving the N_ϕ important feature dimensions. This shows that the GCM's semantic information processing was not as effective as that of SINFONY.

Moreover, unlike in the SNR simulation where a larger number of N_ϕ features always yields better performance, the accuracy does not always increase with the number of features under different expertise levels. For example, the accuracy on the tool dataset with $N_\phi = 10$ and 20 features exceeds the accuracy simulated with 40 features at lower expertise. Even with the highest expertise on the CIFAR10 dataset, the accuracy with 20 features still beats that with 40 features.

This means that GCM is not always capable to learn to effectively extract the semantic information when provided with extra information. This behavior is consistent with the bias-variance trade-off from statistical learning, which explains why low-capacity models generalize better with limited data [43]: GCMs with fewer parameters constrain the hypothesis space of solutions, effectively regularizing the learning process. In contrast, high-capacity GCMs with more inputs and attention weights tend to overfit to a limited knowledge base $\mathcal{D}_{\mathrm{HK}}$ based on a few seen realizations. We conclude that providing more features, i.e., details, to the decision-maker with a small knowledge base $\mathcal{D}_{\mathrm{HK}}$ can lead to a suboptimal decision compared to providing less information.

Main Conclusions: Recalling the design trade-off (20) from Sec. II-F.2, we conclude from both simulation scenario observations that it is more important to match the cognitive capabilities of the GCM by a low-dimensional presentation, i.e, more elaborate SINFONY preprocessing, in this task than providing more relevant information about $\mathbf{z}$ by raw decision features.

However, providing more features instead of the final SINFONY output can have other benefits. For one, we have a slight reduction in processing complexity, since some nodes are removed. Also, in this case, SINFONY is optimized for solving a single task, i.e., deciding whether a tool needs to be replaced or not. In more complex situations, however, HDM may need to deal with unexpected events or changing goals not covered by the current form of our end-to-end sensing-decision framework. A more detailed representation allows the HDM model to react to these changes compared to the probability estimates. Finally, the study focused on decision accuracy alone, using a simulated decision-maker. With human

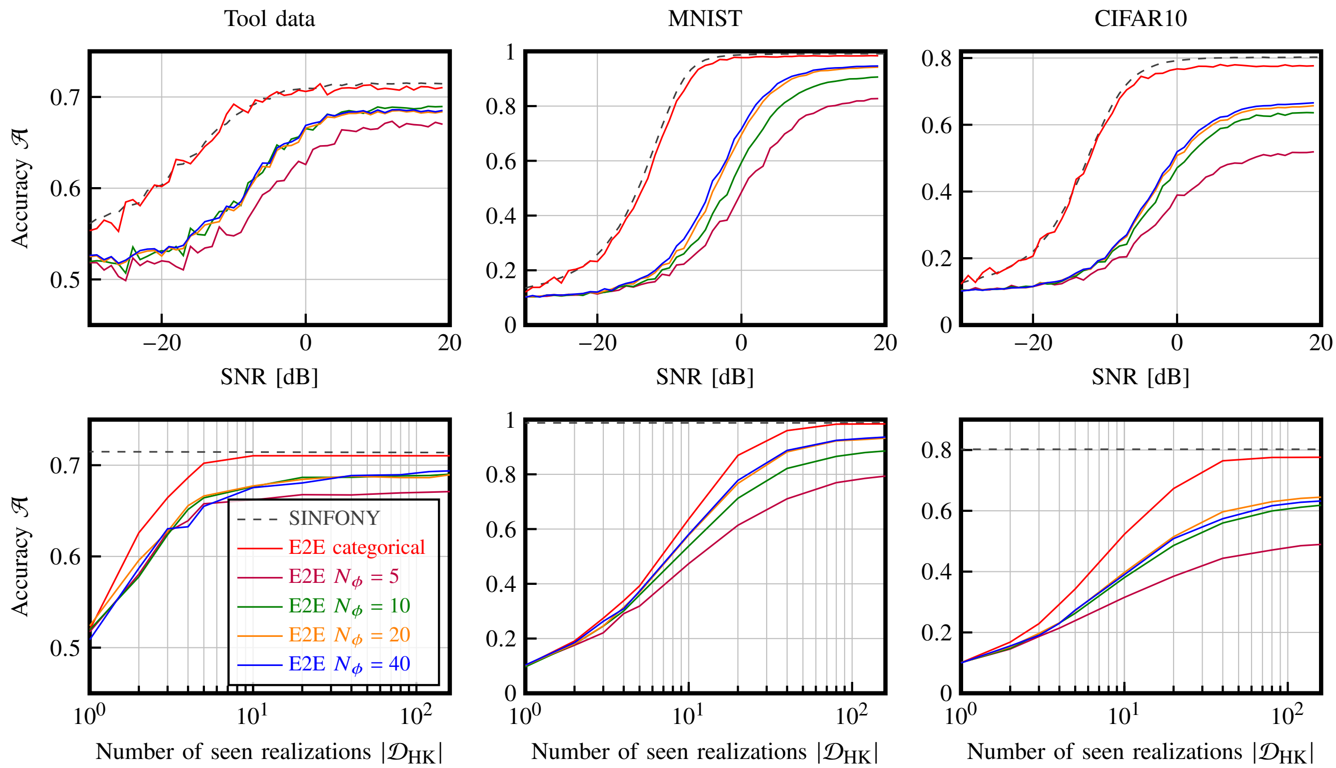

Fig. 4. The simulated performance of the proposed end-to-end sensing-decision framework, including SINFONY and human decisions modeled by the GCM. Each column shows the classification accuracy on different datasets. From left to right: Tools, MNIST, and CIFAR10. The top row shows the accuracy as a function of SNR. The bottom row shows the simulated accuracy as a function of the number of seen realizations. Within each figure, the color of the lines indicates the number of features presented to the GCM.

decision-makers, motivational and emotional aspects such as experienced autonomy and competence will affect performance in addition to information processing ability [74], as we will discuss in the outlook in Sec. IV.

The main takeaways provide answers to the research questions stated at the beginning:

b) *Which information should semantic communication provide for accurate HDM?* The semantic information provided by the SINFONY architecture supports HDM, as motivated by the InfoMax principle. However, using raw decision features leads to imperfect information integration of the HDM model compared to SINFONY, evident through saturation in the simulations. This is consistent with the derived design trade-off and shows that it is more important to match the cognitive capabilities of the GCM by more elaborate SINFONY feature extraction than providing more relevant information.

b) *How much information should semantic communication provide for accurate HDM?* Providing more detailed representations, i.e., more features, does not always increase HDM decision accuracy. The saturation indicates that the HDM model at some point misses subtle details in the additional features. The effect of extra features also depends on the context. For example, with little expertise, more information can misguide instead of help the HDM model which requires carefully balancing the design trade-off, i.e., the information provided by semantic communication with the HDM process.

c) *How does the HDM process impact the end-to-end sensing-decision-making process?* The accuracy of the HDM model can be inferior to that of semantic communication systems due to the probabilistic nature of HDM.

IV. Outlook – Open Questions and Challenges

The proposed end-to-end sensing-decision framework is a first step towards integrating semantic communication and the human receiver. We will now explore remaining open questions and challenges with respect to all our framework components from Fig. 2, and what they mean for semantic communication.

A. Challenge: Optimization of Semantic Communication for Human Decisions

In this article, we have examined how semantic communication affects the decisions of a HDM model in theory and simulations. There remains the question of how semantic communication can be optimized directly for the given human or HDM model to improve decisions.

One idea is to include the human or HDM model in the optimization process. This idea is supported in theory by our extension of the information-theoretic framework on semantic communication via (20) from Sec. II-F, originally aimed to understand both presentation design and HDM model training given a fixed optimized semantic communication system. Including the semantic encoder parameters $\boldsymbol{\theta}$ in maximization of the MILBO (20) allows for joint end-to-end optimization of all components with respect to all framework parameters $\boldsymbol{\theta}, \boldsymbol{\theta}_{\mathrm{P}}, \boldsymbol{\varphi}_{\mathrm{G}}$, leading to a unified design.

However, since both the human and the HDM model are essentially a black box that is not known or differentiable in practice as assumed in this work, this seems difficult. However, it is possible to evaluate the cross-entropy loss or another target metric for the human or HDM model decision and feed it back to SINFONY as a reward. So one idea could be to use the stochastic policy gradient as in [21] to allow optimization over the whole chain, including SINFONY and the human/HDM model.

B. Challenge: Limitations of Human Decision-Making Models

To include the HDM model in an optimization process, it is essential to have an accurate representation of the HDM process. Here, we simulated the decision-making process by applying a computational model, the GCM, for illustrative purposes. While our simulation reflects many traits of HDM with human participants, not all assumptions will apply to realistic categorization decisions. For instance, given the lack of HDM data for the simulated tasks, we assumed perfect memory and optimal performance on the training set for the GCM. These assumptions are difficult to achieve in a real-life situations, but more appropriate assumptions are likely to strongly depend on individuals and tasks.

Accordingly, just like in most of psychological research on HDM, future HDM models will need to be carefully selected and designed to accurately reflect human decision processes for the task of interest, e.g., by accounting for limitations in human information retrieval [53], [75] and contextual influences on the decision process such as limited cognitive resources due to multitasking, acute stress, or time pressure, e.g., [56], [57]. As outlined in Sec. II-D, it is possible to extend the basic GCM (10) used in this work to model many of these HDM aspects. Nevertheless, experimental validation with human participants, typical for psychological experiments, is required to develop and validate appropriate HDM models for the decision problem at hand, assess the beneficial effects of semantic communication and their visualizations, and understand the trade-offs between performance facilitation and potential negative impacts on motivation.

C. Challenge: Presentation of Semantic Information

For tractability in the simulations and to reflect constraints on realistic cognitive processing, we provided the HDM model with two interpretable, handcrafted presentations containing varying amount of information — probability outputs or the most important decision features of semantic communication. The key question is how this abstract representation $\boldsymbol{\nu} = f(\mathbf{y})$ or process $p(\boldsymbol{\nu}|\mathbf{y})$, shown in Fig. 2, translates into real-world scenarios with human subjects. To help humans interpret the output $\mathbf{y}$ or $q_{\boldsymbol{\varphi}}(\mathbf{z}|\mathbf{y})$, it is essential to present it through a Human-Machine Interface (HMI) that connects human users with semantic communication. In practice, this could involve visualizing, e.g., tool damage probabilities, through symbols and colors on a screen or in augmented/virtual reality, with varying levels of detail (see Fig. 1) — from basic machine learning outputs (e.g., tool wear status) to more detailed insights such as algorithm certainty, textual explanations, and even image-based class activation mapping [76].

The HMI is a critical component, as the presentation format can strongly influence the HDM process [77], [78]. It must hence present complex information in a way that enables informed decisions while maintaining essential context. Designing a successful HMI requires an understanding of the domain in which it operates, including the industry, the user base, and the types of decisions being made [79]. A context-aware HMI that adapts to the user's history, preferences, and current situation can further improve decision-making by providing personalized and relevant information [80].

D. Challenge: Variability in Human Decision Goals and Expertise

For many tasks, human decision-makers will differ in their preferences, intentions, and levels of expertise. In addition, task goals may be multi-faceted and subject to change of time, requiring to adjust transmitted information to the individual and the current situation. For example, when communicating information about tools, a person interested in understanding how different machines affect tool usability will need different information than someone focused on identifying tools that need to be replaced. In addition, an expert is likely to prefer a detailed, rich presentation, while a novice may benefit more from clear, concise support.

Furthermore, individual differences can lead to different information even when the decision objective is the same. For example, people differ in how much risk they are willing to accept [81]. Given the same information about the probability that the tool will fail, a risk-seeking person may conclude that the risk is acceptable, while a more risk-averse person would choose to exchange it. Thus, semantic communication that attempts to optimize tool use while keeping the failure rate below a tolerable threshold may require adapting the information conveyed to the decision-maker, for example by changing how potential risks are presented [82], [83]. While the core model (10) of the GCM investigated in this work does not accommodate all individual differences, it can be extended to simulate variability.

E. Challenge: Conflict of Interest between Sender and Receiver

The interests or goals of a human sender may not be well aligned with those of the human receiver. A fundamental factor contributing to such a misalignment of interests could be that human receivers are risk-averse. For example, even if a tool remains functional, the receivers may classify it as defective in order to avoid potential errors, since they are reluctant to take responsibility for using a worn tool. The sender thus has an incentive to manipulate the message in order to influence the receiver's decisions. If the difference in interests is too large, the receiver could ignore any message the sender sends.

This means that successful semantic communication also depends on trust between sender and receiver. Economists, following [84], have long studied this sender-receiver problem using game theory. For a recent overview of this literature, see [85]. They found that the amount of information that can be transmitted depends on how large the difference in interests is. Considering how much the sender wants to manipulate the information to influence the receiver's action is important in semantic communication. Even if the technology allows for very accurate transmission of semantic meaning, the best transmission strategy would still depend on the characteristics of the sender and receiver.

V. Conclusion

In this paper, integrating an interdisciplinary perspective from communications and psychology, we proposed a probabilistic end-to-end sensing-decision framework that wirelessly links sensed data with Human Decision-Making (HDM) by semantic communication. We analyzed this integration exemplarily using SINFONY and an effective HDM model based on generalized context models for specific datasets. The theoretical and numerical results indicate that semantic communication can optimize task performance by balancing information detail with human cognitive processes, achieving accurate decisions while demanding less bandwidth, power, and latency compared to classical methods.

This work is intended to inspire further interdisciplinary research on higher semantic levels of communication. Open questions include how to optimize semantic communication for human decisions, how to extend the HDM model, how to design human-machine interfaces that convey meaning more effectively, and how to account for different intentions between sender and receiver as well as individual differences among receivers.

References

[1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is All you Need," in *31st Conference on Advances in Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, USA, Dec. 2017, pp. 6000–6010.

[2] D. C. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks," in *16th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2013)*. Nagoya, Japan: Springer, Sep. 2013, pp. 411–418.

[3] A. Suresh, E. Beck, A. Dekorsy, P. Rückert, and K. Tracht, "Human-integrated Multi-agent Exploration using Semantic Communication and Extended Reality Simulation," in *10th IEEE International Conference on Automation, Robotics and Applications (ICARA 2024)*, Athens, Greece, May 2024, pp. 419–426.

[4] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, 16th ed. Urbana, IL, USA: The University of Illinois Press, Sep. 1949.

[5] E. C. Strinati and S. Barbarossa, "6G networks: Beyond Shannon towards semantic and goal-oriented communications," *Computer Networks*, vol. 190, p. 107930, May 2021.

[6] X. Luo, H.-H. Chen, and Q. Guo, "Semantic Communications: Overview, Open Issues, and Future Research Directions," *IEEE Trans. Wireless Commun.*, vol. 29, no. 1, pp. 210–219, Feb. 2022.

[7] D. Wheeler and B. Natarajan, "Engineering Semantic Communication: A Survey," *IEEE Access*, vol. 11, pp. 13 965–13 995, Feb. 2023.

[8] D. Gündüz, Z. Qin, I. E. Aguerri, H. S. Dhillon, Z. Yang, A. Yener, K. K. Wong, and C.-B. Chae, "Beyond Transmitting Bits: Context, Semantics, and Task-Oriented Communications," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 5–41, Jan. 2023.

[9] E. Beck, C. Bockelmann, and A. Dekorsy, "Semantic Information Recovery in Wireless Networks," *Sensors*, vol. 23, no. 14, p. 6347, Jul. 2023.

[10] J. Shao, Y. Mao, and J. Zhang, "Learning Task-Oriented Communication for Edge Inference: An Information Bottleneck Approach," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 197–211, Jan. 2022.

[11] R. Carnap and Y. Bar-Hillel, "An Outline of a Theory of Semantic Information," Research Laboratory of Electronics, Massachusetts Institute of Technology, Technical Report 247, Oct. 1952.

[12] F. Zhou, Y. Li, M. Xu, L. Yuan, Q. Wu, R. Q. Hu, and N. Al-Dhahir, "Cognitive Semantic Communication Systems Driven by Knowledge Graph: Principle, Implementation, and Performance Evaluation," *IEEE Trans. Commun.*, vol. 72, no. 1, pp. 193–208, Jan. 2024.

[13] S. R. Pokhrel and J. Choi, "Understand-Before-Talk (UBT): A Semantic Communication Approach to 6G Networks," *IEEE Trans. Veh. Commun.*, vol. 72, no. 3, pp. 3544–3556, Mar. 2023.

[14] J. Bao, P. Basu, M. Dean, C. Partridge, A. Swami, W. Leland, and J. A. Hendler, "Towards a theory of semantic communication," in *IEEE Network Science Workshop (NSW 2011)*, West Point, NY, USA, Jun. 2011, pp. 110–117.

[15] P. Basu, J. Bao, M. Dean, and J. Hendler, "Preserving Quality of Information by Using Semantic Relationships," *Pervasive and Mobile Computing*, vol. 11, pp. 188–202, Apr. 2014.

[16] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep Learning Enabled Semantic Communication Systems," *IEEE Trans. Signal Process.*, vol. 69, pp. 2663–2675, Apr. 2021.

[17] Z. Weng, Z. Qin, and G. Y. Li, "Semantic Communications for Speech Signals," in *IEEE International Conference on Communications (ICC 2021)*, Virtual Conference, Jun. 2021, pp. 1–6.

[18] Z. Weng, Z. Qin, X. Tao, C. Pan, G. Liu, and G. Y. Li, "Deep Learning Enabled Semantic Communications with Speech Recognition and Synthesis," *IEEE Trans. Wireless Commun.*, vol. 22, no. 9, pp. 6227–6240, Sep. 2023.

[19] F. Jiang, Y. Peng, L. Dong, K. Wang, K. Yang, C. Pan, and X. You, "Large AI Model-Based Semantic Communications," *IEEE Wireless Commun.*, vol. 31, no. 3, pp. 68–75, Jun. 2024.

[20] H. Cui, Y. Du, Q. Yang, Y. Shao, and S. C. Liew, "LLMind: Orchestrating AI and IoT with LLM for Complex Task Execution," arXiv preprint: 2312.09007, Aug. 2024.

[21] E. Beck, C. Bockelmann, and A. Dekorsy, "Model-free Reinforcement Learning of Semantic Communication by Stochastic Policy Gradient," in *1st IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN 2024)*, Stockholm, Sweden, May 2024, pp. 367–373.

[22] H. Xie, Z. Qin, X. Tao, and K. B. Letaief, "Task-Oriented Multi-User Semantic Communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2584–2597, Sep. 2022.

[23] X. Luo, R. Gao, H.-H. Chen, S. Chen, Q. Guo, and P. N. Suganthan, "Multimodal and Multiuser Semantic Communications for Channel-Level Information Fusion," *IEEE Wireless Commun.*, vol. 31, no. 2, pp. 117–125, Apr. 2024.

[24] A. Halimi Razlighi, C. Bockelmann, and A. Dekorsy, "Semantic Communication for Cooperative Multi-Task Processing Over Wireless Networks," *IEEE Wireless Commun. Lett.*, vol. 13, no. 10, pp. 2867–2871, Oct. 2024.

[25] A. Halimi Razlighi, M. H. V. Tillmann, E. Beck, C. Bockelmann, and A. Dekorsy, "Cooperative and Collaborative Multi-Task Semantic Communication for Distributed Sources," in *IEEE International Conference on Communications (ICC 2025)*, Montreal, Canada, Jun. 2025, pp. 1–6.

[26] Y. Blau and T. Michaeli, "Rethinking Lossy Compression: The Rate-Distortion-Perception Tradeoff," in *36th International Conference on Machine Learning (ICML 2019)*. Long Beach, CA, USA: PMLR, May 2019, pp. 675–685.

[27] J. Chai, Y. Xiao, G. Shi, and W. Saad, "Rate-Distortion-Perception Theory for Semantic Communication," in *31st IEEE International Conference on Network Protocols (ICNP 2023)*, Reykjavik, Iceland, Oct. 2023, pp. 1–6.

[28] J. W. Payne, J. R. Bettman, E. Coupey, and E. J. Johnson, "A constructive process view of decision making: Multiple strategies in judgment and choice," *Acta Psychologica*, vol. 80, no. 1-3, pp. 107–141, Aug. 1992.

[29] D. Hassabis, D. Kumaran, C. Summerfield, and M. Botvinick, "Neuroscience-inspired Artificial Intelligence," *Neuron*, vol. 95, no. 2, pp. 245–258, Jul. 2017.

[30] D. Kahneman, "A perspective on judgment and choice: Mapping bounded rationality," in *Progress in Psychological Science around the World*, ser. Neural, Cognitive and Developmental Issues. Psychology Press, Jul. 2013, vol. 1, pp. 1–47.

[31] B. Englich, T. Mussweiler, and F. Strack, “Playing Dice With Criminal Sentences: The Influence of Irrelevant Anchors on Experts’ Judicial Decision Making,” *Personality and Social Psychology Bulletin*, vol. 32, no. 2, pp. 188–200, Feb. 2006.

[32] M. Thomas and V. Morwitz, “Heuristics in numerical cognition: Implications for pricing,” in *Handbook of pricing research in marketing*. Cheltenham, UK: Edward Elgar Publishing, Mar. 2008, pp. 132–149.

[33] G. Gigerenzer and W. Gaissmaier, “Heuristic Decision Making,” *Annual Review of Psychology*, vol. 62, no. 1, pp. 451–482, Jan. 2011.

[34] N. Cowan, “The magical number 4 in short-term memory: A reconsideration of mental storage capacity,” *Behavioral and Brain Sciences*, vol. 24, no. 1, pp. 87–114, Feb. 2001.

[35] H. A. Simon, “Bounded Rationality and Organizational Learning,” *Organization Science*, vol. 2, no. 1, pp. 125–134, Feb. 1991.

[36] G. Gigerenzer and D. G. Goldstein, “Betting on One Good Reason: The Take The Best Heuristic,” in *Simple Heuristics That Make Us Smart*, G. Gigerenzer, P. M. Todd, and the ABC Research Group, Eds. New York: Oxford University Press, Apr. 1999, pp. 75–95.

[37] H. Brighton, “Robust inference with simple cognitive models,” in *AAAI spring symposium: Between a rock and a hard place: Cognitive science principles meet AI-hard problems*, vol. 15, Menlo Park, CA, USA, Mar. 2006, pp. 17–22.

[38] G. Gigerenzer and D. G. Goldstein, “Reasoning the fast and frugal way: Models of bounded rationality,” *Psychological Review*, vol. 103, no. 4, pp. 650–669, Oct. 1996.

[39] S. Paul and D. L. Nazareth, “Input information complexity, perceived time pressure, and information processing in GSS-based work groups: An experimental investigation using a decision schema to alleviate information overload conditions,” *Decision Support Systems*, vol. 49, no. 1, pp. 31–40, Apr. 2010.

[40] D. Friedman, “Monty Hall’s three doors: Construction and deconstruction of a choice anomaly,” *The American Economic Review*, vol. 88, no. 4, pp. 933–946, Sep. 1998.

[41] A. N. Sanborn, J.-Q. Zhu, J. Spicer, P. León-Villagrá, L. Castillo, J. K. Falbén, Y.-X. Li, A. Tee, and N. Chater, “Noise in Cognition: Bug or Feature?” *Perspectives on Psychological Science*, vol. 19, no. 1, pp. 123–145, Jan. 2024.

[42] F. I. Seitz, J. B. Jarecki, J. Rieskamp, and B. von Helversen, “Disentangling Perceptual and Process-Related Sources of Behavioral Variability in Categorization,” *PsyArXiv*, Sep. 2024.

[43] O. Simeone, “A Brief Introduction to Machine Learning for Engineers,” *Foundations and Trends® in Signal Processing*, vol. 12, no. 3-4, pp. 200–431, Aug. 2018.

[44] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, “Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion,” *Journal of Machine Learning Research*, vol. 11, no. 110, pp. 3371–3408, Dec. 2010.

[45] O. Simeone, “A Very Brief Introduction to Machine Learning with Applications to Communication Systems,” *IEEE Trans. on Cogn. Commun. Netw.*, vol. 4, no. 4, pp. 648–664, Dec. 2018.

[46] D. J. Koehler and N. Harvey, *Blackwell handbook of judgment and decision making*. Malden, MA, USA: John Wiley & Sons, Apr. 2008.

[47] B. R. Newell, D. A. Lagnado, and D. R. Shanks, *Straight Choices: The Psychology of Decision Making*. London: Psychology Press, May 2022, vol. 3.

[48] J. R. Anderson, “The adaptive nature of human categorization,” *Psychological Review*, vol. 98, no. 3, p. 409, Jul. 1991.

[49] K. J. Kurtz, “Chapter Three - Human Category Learning: Toward a Broader Explanatory Account,” in *Psychology of Learning and Motivation*, B. H. Ross, Ed. Academic Press, Jan. 2015, vol. 63, pp. 77–114.

[50] R. Schlegelmilch, A. J. Wills, and B. von Helversen, “A cognitive category-learning model of rule abstraction, attention learning, and contextual modulation,” *Psychological Review*, vol. 129, no. 6, pp. 1211–1248, Sep. 2021.

[51] R. M. Nosofsky, “Choice, similarity, and the context theory of classification,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 10, no. 1, p. 104, Jan. 1984.

[52] ——, “The generalized context model: an exemplar model of classification,” in *Formal Approaches in Categorization*, A. J. Wills and E. M. Pothos, Eds. Cambridge: Cambridge University Press, Mar. 2011, pp. 18–39.

[53] R. M. Nosofsky and T. J. Palmeri, “An Exemplar-Based Random-Walk Model of Categorization and Recognition,” in *The Oxford Handbook of Computational and Mathematical Psychology*, J. R. Busemeyer, Z. Wang, J. T. Townsend, and A. Eidels, Eds. New York: Oxford University Press, Dec. 2015, pp. 142–164.

[54] S. Bhatia and N. Stewart, “Naturalistic multiattribute choice,” *Cognition*, vol. 179, pp. 71–88, Oct. 2018.

[55] B. J. Meagher and R. M. Nosofsky, “Testing formal cognitive models of classification and old-new recognition in a real-world high-dimensional category domain,” *Cognitive Psychology*, vol. 145, p. 101596, Sep. 2023.

[56] K. Lamberts, “Categorization under time pressure,” *Journal of Experimental Psychology: General*, vol. 124, no. 2, pp. 161–180, Jun. 1995.

[57] F. I. Seitz, B. von Helversen, R. Albrecht, J. Rieskamp, and J. B. Jarecki, “Testing three coping strategies for time pressure in categorizations and similarity judgments,” *Cognition*, vol. 233, p. 105358, Apr. 2023.

[58] J. A. Hoffmann, B. von Helversen, and J. Rieskamp, “Deliberation’s blindsight: How cognitive load can improve judgments,” *Psychological Science*, vol. 24, no. 6, pp. 869–879, Apr. 2013.

[59] D. R. Shanks, “A connectionist account of base-rate biases in categorization,” *Connection Science*, vol. 3, no. 2, pp. 143–162, Jun. 1991.

[60] E. M. Pothos and T. M. Bailey, “The role of similarity in artificial grammar learning,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 26, no. 4, pp. 847–862, Jul. 2000.

[61] D. K. Sewell, T. Ballard, and N. K. Steffens, “Exemplifying “Us”: Integrating social identity theory of leadership with cognitive models of categorization,” *The Leadership Quarterly*, vol. 33, no. 4, p. 101517, Aug. 2022.

[62] T. L. Griffiths and M. L. Kalish, “A multidimensional scaling approach to mental multiplication,” *Memory & Cognition*, vol. 30, no. 1, pp. 97–106, Jan. 2002.

[63] R. M. Nosofsky and T. J. Palmeri, “An exemplar-based random walk model of speeded classification,” *Psychological Review*, vol. 104, no. 2, pp. 266–300, Apr. 1997.

[64] M. D. Lee and D. J. Navarro, “Extending the alcove model of category learning to featural stimulus domains,” *Psychonomic Bulletin & Review*, vol. 9, no. 1, pp. 43–58, Mar. 2002.

[65] R. D. Luce, “On the possible psychophysical laws,” *Psychological Review*, vol. 66, no. 2, p. 81, Jul. 1959.

[66] Y. Zhang, L. Wang, and Q. Wu, “Differential Annealing for Global Optimization,” in *Advances in Swarm Intelligence: Third International Conference on Swarm Intelligence (ICSI 2012)*, ser. Lecture Notes in Computer Science, vol. 7331. Berlin, Heidelberg: Springer, Jun. 2012, pp. 382–389.

[67] T. Serre, “Models of Visual Categorization,” *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 7, no. 3, pp. 197–213, May 2016.

[68] K. N. Prasad and B. Ramamoorthy, “Tool Wear Evaluation by Stereo Vision and Prediction by Artificial Neural Network,” *Journal of Materials Processing Technology*, vol. 112, no. 1, pp. 43–52, May 2001.

[69] B. Papenberg, S. Hogreve, T. Bochardt, C. Bornholdt, T. Heinrich, and K. Tracht, “Influence of Illumination on the Image-Based Classification Accuracy of Wear on Milling Tools,” in *17th CIRP Conference on Intelligent Computation in Manufacturing Engineering (ICME 23)*, Ischia, Italy, Jul. 2023, pp. 366–371.

[70] E. Beck, “Semantic Information Transmission and Recovery (SINFONY) Software,” Zenodo, Dec. 2024. [Online]. Available: https://doi.org/10.5281/zenodo.8006567

[71] K. He, X. Zhang, S. Ren, and J. Sun, “Identity Mappings in Deep Residual Networks,” in *14th European Conference on Computer Vision (ECCV 2016)*, ser. Lecture Notes in Computer Science, Amsterdam, Netherlands, Oct. 2016, pp. 630–645.

[72] ——, “Deep Residual Learning for Image Recognition,” in *29th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[73] R. M. Battleday, J. C. Peterson, and T. L. Griffiths, “Capturing human categorization of natural images by combining deep networks and cognitive models,” *Nature Communications*, vol. 11, no. 1, p. 5418, Oct. 2020.

[74] R. M. Ryan and E. L. Deci, “Self-Regulation and the Problem of Human Autonomy: Does Psychology Need Choice, Self-Determination, and Will?” *Journal of Personality*, vol. 74, no. 6, pp. 1557–1586, Dec. 2006.

[75] N. Stewart, N. Chater, and G. D. A. Brown, “Decision by sampling,” *Cognitive Psychology*, vol. 53, no. 1, pp. 1–26, Aug. 2006.

[76] S. Y. Lee, B. A. Tama, S. J. Moon, and S. Lee, “Steel Surface Defect Diagnostics Using Deep Convolutional Neural Network and Class Activation Map,” *Applied Sciences*, vol. 9, no. 24, p. 5449, Dec. 2019.

[77] S. T. Hawley, B. Zikmund-Fisher, P. Ubel, A. Jancovic, T. Lucas, and A. Fagerlin, “The impact of the format of graphical presentation on health-related knowledge and treatment choices,” *Patient Education and Counseling*, vol. 73, no. 3, pp. 448–455, Nov. 2008.

[78] A. Strathie, G. Netto, G. H. Walker, and G. Pender, “How presentation format affects the interpretation of probabilistic flood risk information,” *Journal of Flood Risk Management*, vol. 10, no. 1, pp. 87–96, Mar. 2017.

[79] K. Nikolova, “Framework and Practical Guidelines for Creating a Current State Analysis for (Re)Designing an HMI within the Industry 4.0 Era: Case Study,” Master’s thesis, University of Twente, Enschede, Netherlands, Aug. 2022.

[80] S. Meftah, M. Sahnoun, M. Messaadia, and S. M. Benslimane, “Context Aware Human Machine Interface for Decision Support,” in *International Conference on Cyber Management and Engineering (CyMaEn 2023)*, Bangkok, Thailand, Jan. 2023, pp. 143–147.

[81] C. Wang, S. M. O’Neill, N. Rothrock, R. Gramling, A. Sen, L. S. Acheson, W. S. Rubinstein, D. E. Nease, and M. T. Ruffin, “Comparison of risk perceptions and beliefs across common chronic diseases,” *Preventive Medicine*, vol. 48, no. 2, pp. 197–202, Feb. 2009.

[82] C. Franco, J. L. Hougaard, and K. Nielsen, “Handling Risk Attitudes for Preference Learning and Intelligent Decision Support,” in *Modeling Decisions for Artificial Intelligence*, ser. Lecture Notes in Computer Science, V. Torra and T. Narukawa, Eds. Springer International Publishing, Sep. 2015, vol. 9321, pp. 78–89.

[83] J. Liesiö, M. Kallio, and N. Argyris, “Incomplete Risk-Preference Information in Portfolio Decision Analysis,” *European Journal of Operational Research*, vol. 304, no. 3, pp. 1084–1098, Feb. 2023.

[84] V. P. Crawford and J. Sobel, “Strategic Information Transmission,” *Econometrica*, vol. 50, no. 6, pp. 1431–1451, Nov. 1982.

[85] A. Blume, E. K. Lai, and W. Lim, *Chapter 13 Strategic information transmission: a survey of experiments and theoretical foundations*. Cheltenham, UK: Edward Elgar Publishing, Oct. 2020.

Edgar Beck (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Bremen, Germany, in 2014 and 2017, respectively, where he is currently pursuing a Ph.D. degree in electrical engineering at the Department of Communications Engineering (ANT). His research interests include cognitive radio, compressive sensing, massive MIMO systems, semantic communication, and machine learning for wireless communications. He was a recipient of the OHB Award for the best M.Sc. degree in Electrical Engineering and Information Technology at the University of Bremen in 2017.

Hsuan-Yu Lin is a postdoctoral researcher at the Department of Psychology at the University of Bremen. He received his M.Sc. degree in psychology from National Chengchi University in Taiwan and his Ph.D. in psychology from University of Zürich in Switzerland. His research focus is on decision-making processes, particularly patch-leaving and sampling decisions.

Patrick Rückert is chief engineer at the Bremen Institute for Mechanical Engineering at the University of Bremen. He received his B.Sc. in industrial engineering from the University of Applied Sciences in Karlsruhe and his M.Sc. in production engineering at the University of Bremen. His research focuses on human-robot collaboration, augmented and virtual reality systems, and the integration of machine learning for assembly processes.

Yongping Bao is a postdoctoral researcher at the SOCIUM institute of the University of Bremen. She received her Ph.D. in Economics from the University of Konstanz. Her research focuses on behavioral and experimental economics, game theory, and human-algorithm interaction.

Bettina von Helversen is a professor for Experimental Psychology at the University of Bremen since 2019. Before, she held a Swiss Science Foundation Professorship at the University of Zurich for Cognitive Decision Psychology and worked at the Center for Economic Psychology in Basel. She received her Ph.D. from the Humboldt University in Berlin. She is interested in understanding and modeling how people make judgments and decisions. In her work she focuses on the different cognitive strategies people use to solve these tasks and the factors that influence strategy selection such as task structure, memory, affect or stress and how these change over the life span.

Sebastian Fehrler joined the University of Bremen, SOCIUM, as Professor of Economics in 2020. He studied Statistics (B.Sc., HU Berlin) and Economics (M.Sc., U Nottingham; Ph.D., U Zurich), spent a year as a visiting postdoc at the Center for Experimental Social Sciences at NYU, and worked as an assistant professor at the University of Konstanz. His research interests are in the fields of public, organizational and behavioral economics. He uses game theory and experiments to address his research questions.

Kirsten Tracht is professor at the faculty of production engineering of the University of Bremen and speaker of the Bremen Institute for Mechanical Engineering. She studied mechanical engineering at the University of Hannover and received her Ph.D. under the supervision of Prof. H.K. Tönshoff in the field of die and mold making. She received the Kienzle-Medal. She is responsible for the master’s study course ‘production engineering’. Her research focuses on assembly technologies and production design techniques in industrial settings.

Armin Dekorsy (Senior Member, IEEE) is a professor at the University of Bremen, where he is the director of the Gauss-Olbers Space Technology Transfer Center and heads the Department of Communications Engineering. With over eleven years of industry experience, including distinguished research positions such as DMTS at Bell Labs and Research Coordinator Europe at Qualcomm, he has actively participated in more than 65 international research projects, with leadership roles in 17 of them. He is a Senior Member of the IEEE Communications and Signal Processing Society and a member of the VDE/ITG Expert Committee on Information and System Theory. He co-authored the textbook 'Nachrichtenübertragung, Release 6, Springer Vieweg', which is a bestseller in the field of communication technologies in German-speaking countries. His research focuses on signal processing and wireless communications for 5G/6G, industrial radio, and 3D networks.