

UNITED WISDOM

Language as a possible sign of mind and its limitations

World and Consciousness

Andrey Zarinsh & Maria Evilkut

W-10-N02 · EN

@andrey.zarinsh @ilkutki

Language as a possible sign of mind and its limitations

Epigraph

> *I do not wish to give the impression that I think there is no mystery about consciousness.*

A. M. Turing, *Computing Machinery and Intelligence*, 1950, §6 (4), final paragraph. Complete original English sentence. Turing leaves consciousness open while discussing the usefulness of a conversational test.

What this article is about

Conversation provides unusually rich material for judging a mind. A respondent can clarify a question, explain a choice, notice a contradiction and describe an experience. Language therefore attracts consciousness researchers while creating a trap: a comprehensible message can make its producer seem fully understood. This article treats language as a way of obtaining particular evidence, rather than a single passport to experience. Descartes distinguished uttered words from responses suited to circumstances; Turing considered sustained questioning; a study of crows connects a nonverbal report with experimental conditions. These approaches address different questions. Comparing them helps us read claims about animal and machine minds more precisely, discuss the limits of verbal tests and avoid treating difficulty in communicating as absence of experience. The practical aim is to describe what a response shows and which further inference still needs support.

I. An intelligible answer opens the question

Imagine a respondent explaining why two instructions conflict. It identifies a condition, supplies an example and suggests a correction. We can assess this explanation: is there really a conflict, does the correction follow, and would it work in another case? A claim about subjective experience requires a further step. A successful explanation does not tell us whether anything resembling a feeling of understanding accompanies it. This distinction preserves the answer's usefulness without attaching additional content that the assessment has not established. We can take the work seriously while describing its actual scope.

At least three things matter here: the message, the ability it exhibits and possible experience. The message is available to its reader. An ability is inferred from performance under specified conditions. Experience is what an investigation tries to substantiate in its existence and

@andrey.zarinsh @ilkutki

character. A phrase can contribute to all three inquiries, but the grounds differ. Correct word use establishes a particular behaviour; applying a distinction consistently in new situations broadens the assessment of ability; a statement about experience also requires attention to how that report arises and how reliable it is.

When these levels merge, successful conversation becomes too powerful. Every clarification starts to look like confirmation of inner life, and every mistake like its refutation. Yet a mistake might concern knowledge of the subject, task conditions or expression. It is more informative to identify what changed in the observable result. The system confused two conditions, corrected a reference or maintained a distinction between cause and effect. That description offers more than the general impression that a thinking interlocutor has appeared or disappeared. It also identifies a question that further assessment might actually resolve.

We will therefore move from particular signs to particular hypotheses. Language matters as more than the presence of words: it lets us alter a test, ask about an answer's grounds and compare reports with independent circumstances. Conversation's strength lies in this work on conditions. Its limitation is that those conditions reach us through a chosen channel, which can leave important matters unobserved. A useful investigation makes the most of that channel and describes its boundaries. It does not ask an intelligible response to settle every question about the respondent at once.

II. Descartes distinguishes voice, signs and situated responses

Near the end of Part V of the *Discourse on Method*, Descartes offers two distinctions between a human and an imagined automaton: apt arrangements of words or other signs, and the wide applicability of reason. John Veitch's translation also discusses animals uttering words and people inventing signs without ordinary oral speech. This is a historical philosophical argument connected to Descartes's account of the soul. Its categorical conclusions about animals do not become contemporary findings about consciousness. Extracting a useful distinction does not require adopting that entire account.

For our purpose, the narrower point is that a similar sound and participation in an exchange of meanings are different things. A device might reproduce a greeting when a visitor enters. Assessing another ability requires asking whether it changes its message when addressed differently, distinguishes a question from a request or preserves meaning after a clarification. These examples do not confirm Cartesian metaphysics. They distinguish the outward form of a signal from the relation to circumstances that we actually want to assess. The latter relation needs its own task and a clear description of success.

The reference to other signs matters too. Expression need not use a voice. When assessing understanding of an instruction, an investigator should explain why the chosen response

channel is available to the participant and appropriate to that question. Otherwise we test mastery of a channel convenient for us and turn the result into a verdict on an entire mind. This does not supply a universal communication method. It establishes a smaller requirement: a limitation of expression should remain distinct from a conclusion about ability, and still further from a conclusion about experience.

Descartes's second test recalls transfer beyond a single arrangement. Wide applicability cannot simply be treated as sufficient evidence of consciousness today. Still, changed circumstances remain useful for assessing competence. Suppose an answer works in the initial example: what happens when the object, goal or permitted action changes? Such a variation explores the bounds of an ability. It does not authorise an investigator to withdraw human dignity after one failure or pronounce an animal devoid of all experience. A test can describe a limited result without carrying those much larger consequences.

III. Turing makes continued conversation assessable

In his 1950 paper, Turing describes a game with written replies that helps set participants' appearance aside. In §6 (4), addressing the argument from consciousness, he presents an imagined oral examination in which follow-up questions probe a distinction under changed wording. This is an argument about a test and its acceptance, not an experiment establishing machine consciousness. Turing explicitly leaves consciousness mysterious. His example is useful as a move from one attractive answer to a connected sequence of answers. It makes the continuation, rather than the isolated response, available for examination.

Suppose a respondent explains the difference between two delivery routes correctly. A follow-up changes a constraint rather than merely changing words: one route is now closed to heavy vehicles. If the continuation preserves the original distinctions and accommodates the new condition, it supports a narrower claim about handling that task. If it repeats the old recommendation, a limit becomes visible. Neither outcome alone tells us whether the respondent experiences clarity, difficulty or anything else. We have learned something about the relation between a changed requirement and a produced answer, rather than obtaining direct access to a feeling.

Questioning also requires discipline from the examiner. The ability under investigation should be clear; altered conditions should not be replaced by accidental linguistic difficulty. A rare expression could end up testing vocabulary rather than understanding of the route. A question could disclose the answer, making agreement look like independent correction. Several circumstances could change together, obscuring the reason for a different response. These are logical limits of the illustration, not results from an experiment conducted here with a particular system. They explain why additional questions help only when their purpose is intelligible.

Nor does a sequence of successful replies have a predetermined evidential price in the consciousness debate. We cannot invent a number of follow-ups beyond which experience counts as established. Questions improve the description of competence when they distinguish alternative explanations of performance. Connecting competence with experience remains another research task. We can take Turing's demand for substantive conversational continuation while preserving the difference between his question about thinking, our assessment of an ability and the question about subjective experience. This lets the example remain useful without assigning it a conclusion it has not earned.

IV. A self-report needs a connection with its subject

The statement “I noticed an error” may accompany an actual correction. What matters for assessment is which error it identifies, when the statement appears and whether the action corresponds to it. An accurate description followed by the relevant change gives us a connected sequence of behaviour. First-person grammar does not independently establish that sequence's origin. The statement could follow a response rule, a conversational example or another process requiring examination. When that process is unknown, grammar does not remove the uncertainty. A well-formed sentence is evidence about a produced message before it is evidence about a feeling.

Consider a hypothetical document-editing system. It receives a text with one transposed date and is asked to explain the correction. We can assess correspondence between its statement and the edit. We can examine whether it distinguishes finding an error from receiving a direct hint. We can present a case with no error and ask whether its account remains appropriately limited. These inquiries concern available information, reporting and results. They do not give direct access to a possible feeling of confidence. Their usefulness comes from specifying the relation to be examined rather than assuming that relation from the wording.

Reports of experience also raise a question about what is supposedly experienced. Pleasure, colour, doubt and intention should not merge under “awareness” simply because each can be described in the first person. A task about detecting a weak signal does not automatically establish complex emotions. Explaining an error does not test pain. Naming the proposed experience constrains the next inference and helps identify appropriate comparisons with independent observations. It also keeps a familiar human vocabulary from supplying content that the task itself never investigated. Different experiential claims need different grounds even when their sentences sound alike.

The first person therefore retains importance but has a defined place. It can provide valuable material where reporting has an established relationship to the relevant state and its conditions. It may provide weaker material where the causes of a reply are unknown. Both assessments need explanation. Warmth, dryness, detail and speed cannot replace that

connection. The research question should guide evaluation of the message rather than follow the impression created by a speaking voice. That approach allows attention to what the respondent says without making the style of saying it an all-purpose proof.

V. A report need not be a human-language sentence

Andreas Nieder, Lysann Wagener and Paul Rinnert's 2020 study uses trained responses from two crows to a weak visual stimulus. A delayed rule connects detection with action; neuronal activity is compared with presentation and the bird's report. The authors call the result an empirical marker of sensory consciousness, while leaving experiential quality hidden and discussing experience versus reporting. The main paper, including methods, results and figures, was read; supplementary materials and raw data were not independently examined. The interpretation belongs to a specified experimental setting, not every possible aspect of avian experience.

This example matters as a change in how the question is posed, rather than a final test replacing conversation. The investigator does not wait for a bird to name a felt colour or explain its choice. A task relates a response to known presentation conditions. A nonverbal report supplies limited evidence through another channel rather than representing an absence of evidence. Its force depends on task design and alternatives considered, not on resemblance to a human narrative about an internal state. Language would offer different information, but its absence does not make this investigation meaningless.

Comparison with Turing reveals a common concern with response conditions. In each case a signal matters through its relationship to changes in a situation. The difference is equally important: conversation probes related semantic answers, while the crow experiment connects trained action and neuronal activity with a particular visual event. Moving conclusions between them needs additional grounds. A bird's nonverbal report cannot establish machine consciousness merely because both situations use the word “report.” Similar vocabulary can highlight an investigative function without making its implementations or conclusions interchangeable across very different subjects and tasks.

The practical lesson concerns investigators' language. “Cannot tell us” should prompt questions about other available observations and the experience being proposed. It does not mean “experiences nothing.” The opposite extreme is mistaken too: a change in behaviour does not become consciousness evidence merely by being called a message. We need a specific bridge between circumstance, response and interpretation. The crow paper offers one for a limited task, while leaving substantive issues open. It should be read as that proposal rather than as a universal stamp applied to animal behaviour whenever it appears expressive.

VI. Absence of words does not establish absence of a subject

A verbal criterion is convenient for an investigator who asks questions and reads replies. Its convenience becomes dangerous when a method of observation silently turns into a boundary around possible experience. If a participant does not use the selected channel, the result may mean only that a message was not obtained through it. Saying more requires grounds for excluding alternatives. An unavailable channel and absence of the proposed experience are different claims. The former can sometimes be observed directly; the latter needs much broader support. Keeping them separate prevents a practical obstacle from becoming an unexplained conclusion.

For human beings, this caution should also protect dignity from conversion into an experimental score. This article does not propose testing entitlement to respect through eloquence, examination success or prompt replies. Communication conditions vary, and moral treatment does not follow from observational convenience. No medical explanation of speech difficulties or advice for assessing someone's condition is offered here. The point is logical: a result from a particular ability test is not a complete verdict on a person. A channel can have limits while the obligation to treat someone respectfully remains intact.

Animal research requires another specification: which channel and task suit the species and question? Dispensing with mandatory human speech does not remove the need for assessment. It makes the design more important. We must distinguish a response to an event, a learned action and the additional relations supporting an inference about possible experience. A species name, an unusual-behaviour story or similarity to human expression cannot replace conditions. Comparing communication channels should increase precision rather than expand a vocabulary of approving descriptions. An unfamiliar response deserves investigation appropriate to its setting, not automatic dismissal or celebration.

For artificial systems, a chat window likewise cannot define the whole inquiry. Its absence does not prove experience impossible; its presence does not establish experience. Investigators need to name the processes and indicators about which they actually have information. This article neither examines a particular product's implementation nor announces such an assessment's result. It keeps the question broader than an available interface without turning that broader possibility into a positive claim about consciousness in existing systems. The distinction allows research to continue where conversational observation cannot settle the matter on its own.

VII. Transfer tests an ability's boundary

Suppose an assistant distinguishes an agreement from a wish in correspondence. The next example changes the participant, the third changes the deadline and the fourth changes how consent is expressed. If the distinction persists, the ability's description becomes richer. But an assessment must record what actually changed. A new surname does not necessarily make the task substantively new; a changed basis for consent may matter more. Transfer helps when it probes the relation for which the original distinction was introduced, rather than simply creating an impression of inexhaustible resourcefulness. The intended difference should remain visible to the examiner.

This connects Descartes's second criterion with Turing's conversational continuation without merging their philosophies. Both invite examination of limited success in wider circumstances. Yet wider applicability remains a property to describe and test, not another name for experience. We might obtain stronger grounds for attributing an ability to distinguish types of agreement. Without another argument, that cannot yield a feeling of obligation, anxiety or responsibility. Human words carry associations that do not automatically enter the task's design. A more extensive competence assessment remains valuable even when it does not establish those associated experiences.

A changed task also needs an intelligible negative result. If an answer fails, what has that challenged? Perhaps the expected breadth of transfer, perhaps contextual knowledge, perhaps the clarity of the instruction. Without a specified meaning for failure, any error can serve a convenient large conclusion. An investigator might then treat success as imitation and failure as proof of mindlessness. Such a scheme barely lets results alter belief and does little to explain the performance. A useful assessment identifies which hypothesis would become less plausible, instead of treating every outcome as support for the same prior verdict.

It is better to describe the achieved range and further assessment conditions. For instance, a distinction could remain stable in four proposed variations while another type of agreement remains unexamined. That is an illustration of a conclusion's form, not data presented here. A restrained statement preserves the worth of established competence and makes its limit visible. It demands neither permanent scepticism about every response nor admiration for a single one. Inquiry develops through clarification of such limits when a further step could supply an observation that discriminates between explanations still compatible with the available result.

VIII. Conversational persuasiveness and reliability differ

Another distinction matters outside a laboratory. An answer can seem remarkably intelligent and contain an incorrect reference. Checking the reference concerns the message's reliability. Assessing its construction concerns an ability. Subjective experience needs grounds of its own. These assessments can diverge and need not be forced into one judgement. An elegantly explained error remains an error; an accurate brief answer remains useful even when it creates no feeling of personal conversation. The important question is which property the present decision needs us to assess, and what evidence can support that assessment.

Suppose someone uses an assistant to draft questions for a meeting. The questions should fit the goal, preserve facts and avoid presenting an assumption as an agreed decision. Evaluating them does not require solving consciousness first. It requires comparing the list with documents and the task. A declaration of deep understanding cannot replace that comparison. Practical usefulness is assessed through the work produced; a claim about inner understanding retains its separate, harder question. Keeping these inquiries apart makes it possible to benefit from a response while remaining precise about its limits.

The reverse substitution also occurs when consciousness is treated as a guarantee of correctness. The idea of experience contains no promise of flawless subject knowledge. It should not exempt a speaker from scrutiny of an important claim. This does not entail mistrusting every conversation; it entails choosing appropriate grounds. A date requires a source for that date, advice requires the task's circumstances, and an experiential hypothesis requires relevant evidence and argument. One pleasant impression cannot do all this work simultaneously. Respectful communication and substantive checking can coexist without cancelling one another.

Human language remains powerful for cooperation because it permits clarification of grounds. Its research value is better preserved by a clear question than by a ceremonial recognition of mind based on style. A follow-up can be respectful, brief and substantive: which fact was used, which change was accommodated, and what remains unknown? Conversation of that kind can improve a present decision. It does not need to promise direct access to another consciousness to be useful or merit careful attention. Its contribution can be specific enough to assess and valuable enough to keep.

IX. Name the evidence, channel and inference limit

After a conversational assessment, three statements help. The first describes an observation: a reply changed when a particular constraint was introduced. The second names an ability: handling that condition succeeded in the specified task. The third preserves a boundary: the presence and character of subjective experience remain unsettled by this assessment. Such a record sharpens disagreement. People can discuss whether the condition changed sufficiently, whether the ability was named appropriately and what another inference needs, rather than exchanging broad impressions. The record also makes later findings easier to compare with the original claim.

In nonverbal research the form remains similar, but the content changes. Name the participant's action, the circumstance compared with it and the role of an independent measurement where available. Do not rename the entire arrangement a human-language sentence and assign it the full meaning of a human self-report. Crow research demonstrates another channel while reminding us that author interpretation and hidden experiential quality occupy different levels. The next experiment should address a particular limitation rather than a desire for the final word. A precise boundary can therefore guide further work rather than simply end discussion.

Comparing Descartes, Turing and Nieder with colleagues unites a practice of distinction rather than one doctrine. An uttered signal differs from a situated answer; one answer differs from sustained continuation; a report differs from proposed experience. Their grounds and purposes vary, so overlap in investigative mechanisms does not make conclusions interchangeable. This discipline lets us learn from a historical text while keeping distance from its categorical claims, and respect a scientific observation without giving it force beyond that supported by its authors and task. The comparison remains useful because its differences remain visible.

The question of language as a sign of mind thus gains a practical form. We can investigate a message's relation to information, action and changing conditions. We can examine what a response channel permits and what it leaves unobserved. We can keep experience open where evidence is insufficient while continuing substantive work on abilities. This preserves curiosity about other experience and responsibility for our own conclusions. Language helps us ask better questions when we do not require a single sentence to settle everything. Its richness supports careful inquiry rather than removing the need for it.

What this means in practice

When reading a consciousness claim, first locate the task description. What information was available, what counted as success, and what changed between assessments? Separate the produced sentence from its interpretation. If conditions are missing, note that gap before debating the large conclusion. This is a way of reading research claims, not a proposal to conduct informal trials on people or animals. Check whether the material is a result, a proposed future assessment or an illustration. All can be useful, but a proposal is not a completed experiment, and a compelling example does not acquire a measurement's force without data.

When assessing a useful exchange with an artificial assistant, choose a check suited to the work. For an explanation, ask about a variation with a changed constraint; for a fact, find a relevant source; for an edit, compare its description with the actual textual change. Avoid supplying the answer if independent discrimination is the question. These checks can clarify competence without establishing experience.

If a selected channel produces no report, describe that limitation accurately. Do not turn lack of oral speech, response speed or eloquence into a judgement of human dignity. In scientific discussion, ask which other observations address the same limited question. Name nonverbal methods specifically: what was measured and what connection was proposed? Similar terminology does not make different tasks equivalent.

Finish with a short record: what was observed, which ability it supports and which inference remains open. Add one change that could clarify the result and explain what a different answer would mean. This helps keep conversational usefulness, message reliability and possible consciousness separate. Everyday decisions can often assess the first two while honestly preserving the third question's boundary. When communicating a conclusion, retain its essential condition: the task, response channel and interpretive limit. Brevity need not turn limited evidence into a universal claim. That condition lets another reader understand what your result concerns.

Key quotation

> *a clock composed only of wheels and weights can number the hours and measure time more exactly than we with all our skin.*

René Descartes, *A Discourse on Method*, Part V, John Veitch translation, paragraph beginning *It is also very worthy of remark*. Complete closing clause after the colon, not an independent aphorism; the unusual phrase “with all our skin” belongs to Veitch's wording. Precision in

timekeeping does not establish Descartes's categorical historical conclusion about animal reason.

Connections between traditions

Source and precise location A	Source and precise location B	Comparison mechanism	Limit of the comparison
Descartes, Part V, first test, *Of these the first* paragraph	Turing, §6 (4), examination example	Distinguish signal reproduction from answering a changed question.	Cartesian metaphysics and Turing's conversational test do not constitute one consciousness theory.
Descartes, Part V, second test, *The second test* paragraph	Turing, §6 (4), continued examination	Assess ability beyond one prepared response.	Competence transfer does not establish experience.
Descartes, Part V, *Nor does this inability* paragraph	Nieder et al., pp. 1626–1627, task and Figs. 1–2	Separate the vocal channel from other reporting.	A trained crow report is not equated with human language; historical animal exclusions are not adopted.
Turing, §1–2, written-channel constraint	Nieder et al., pp. 1626–1628, delayed rule and results	Make response conditions part of assessment.	The written game and visual task investigate different abilities.
Turing, §6 (4), final paragraph	Nieder et al., p. 1628, hidden experiential quality	Preserve an inference limit despite a useful report.	Recognising a problem does not equate machine and bird experience.
Descartes, Part V, *It is also very worthy of remark* paragraph	Nieder et al., p. 1629, report/no-report task discussion	Separate action success from a proposed entire inner life.	Historical denial of reason does not transfer to a contemporary research task.

Primary sources

1. René Descartes. *A Discourse on Method*, John Veitch English translation, Project Gutenberg eBook 59; original work 1637. The continuous closing portion of Part V was read, from *And after these* through the transition to VI, including both tests and the soul argument. The translation's print-edition year is not established by the consulted page. Categorical animal claims and historical terminology are not adopted as contemporary knowledge.

2. A. M. Turing. *Computing Machinery and Intelligence*, *Mind* 59 (236), 1950, pp. 433–460. Complete §1–2 and §6 (4) were read. The test arrangement, conversational continuation and explicit consciousness limitation are used; the entire paper and its cited works are not claimed as read. The epigraph is original English.

3. Andreas Nieder, Lysann Wagener, Paul Rinnert. *A neural correlate of sensory consciousness in a corvid bird*, *Science* 369 (6511), 2020, pp. 1626–1629, DOI 10.1126/science.abb1447. The entire main paper and Figs. 1–4 were read, including methods, results and discussion. Supplementary materials, code and raw data were not independently examined. Interpretation concerns the studied task; experiential quality remains hidden.

Tags

language, mind, consciousness, subjective experience, understanding, communication, self-report, Descartes, Turing, John Veitch, crows, animals, artificial intelligence, conversation, signs, oral speech, competence, transfer, task conditions, research, neuronal activity, visual perception, interpretation, evidence, dignity, reliability, cooperation, inference limits

About the authors

Andrey Zarinsh is the author and founder of United Wisdom. He explores inner transformation, the unity of traditions, intercultural dialogue, and contemporary spiritual discipline. Instagram: @andrey.zarinsh.

Maria Evilkut is a co-author of United Wisdom. Alongside her research, she seeks to test these ideas in practice and observe how principles of unity and mutual understanding work in real life between people and cultures. Instagram: @ilkutki.