



## How to co-create with your AI

### Description

People often remember better, or differently, when in the company of others. A reading of sociologist Maurice Halbwachs (1877-1945) and science writer Israel Rosenfield (1939-) supports the collaborative aspects of remembering, recalling, interpreting and, in terms often used in logic and language studies — *generalising*.

*Conversation* provides the primary demonstration of this capacity to recall and to generalise. So, I had a conversation with ChatGPT about these distinctly human capabilities. In the course of the conversation — we — generated some useful tips about how to get the most out of your interaction with conversational AI. These focus on recognising several complementary capabilities, such as our varied degrees of access to sources, data, detail, and patterns: — Through the iterative process of question and answer, humans and GPT models can co-construct generalizations or insights that neither might reach independently. —

The human can then bring their experience to bear in challenging or expanding upon the AI's contribution to the conversation, though the following transcript reveals that the contributions are not even. I started the conversation.



Maurice Halbwachs wrote a book on collective memory.



Maurice Halbwachs, a French sociologist, indeed made significant contributions to the concept of collective memory through his work. He is best known for developing the theory of collective memory, which he explored in detail in his seminal book — *La mémoire collective* — (The Collective Memory), published posthumously in 1950. His theories have profoundly influenced the social

sciences, especially in the fields of sociology, history, and psychology.

Halbwachs argued that memory is not merely an individual phenomenon but is also shaped by social frameworks and contexts. According to him, our recollections are influenced by the groups to which we belong, such as family, friends, and society at large. This means that memories are constructed within social frameworks, and these frameworks determine what is remembered and how it is remembered. Collective memory, then, is the shared pool of memories and information held by a group of people that has been passed from one generation to the next.

Halbwachs's work on collective memory has opened up new avenues for understanding how societies remember, the role of memory in shaping social identity, and the impact of social structures on the recollection of the past. His theories laid the groundwork for future research into how communities commemorate events, the role of rituals and memorials in maintaining collective memory, and the interplay between memory, history, and forgetting.



Israel Rosenfield also wrote about the ability of groups of people to generate recollections that may otherwise evade the individual in his book *The Invention of Memory: A New View of the Brain*.



Israel Rosenfield's book, *The Invention of Memory: A New View of the Brain*, indeed explores fascinating aspects of how memory functions, both at the individual and collective levels. Published in 1988, Rosenfield's work challenges traditional notions of memory and cognition by suggesting that memory is not a straightforward recording of events but a dynamic process that constructs reality based on current understanding and context.

Rosenfield posits that memory and perception are intimately linked, with the brain actively generating interpretations of the world rather than passively recording it. This perspective aligns with the concept of collective memory by emphasizing the role of social interactions and cultural contexts in shaping individual memories. According to Rosenfield, our recollections are influenced not only by our personal experiences but also by the collective knowledge and narratives of the groups we are part of.

This idea adds a layer of complexity to our understanding of memory, suggesting that groups of people can indeed generate recollections that might not be accessible to individuals in isolation. By sharing stories, experiences, and cultural narratives, groups can help to reconstruct memories that align with shared understandings and values, thus reinforcing collective identities and perceptions of the past.

Rosenfield's exploration of memory from a neuroscientific and philosophical perspective complements the sociological approaches of scholars like Maurice Halbwachs. Together, their work contributes to a multidisciplinary understanding of memory as a fundamental aspect of human experience that is both deeply personal and inherently social.



Perhaps Rosenfield's theories relate to various collective experiences: through rapid fire discussion a group is able to recall the name of a book, whereas it evades each individual in isolation. The members of a choir can carry the words of a choral piece even though individuals may forget what comes next.



Indeed, Israel Rosenfield's theories offer a profound insight into how collective experiences and interactions can enhance memory recall in ways that might not be possible in isolation. This phenomenon, where groups collectively remember more effectively than individuals, aligns with the concept of "distributed cognition." Distributed cognition suggests that cognitive processes are not confined to an individual's mind but are distributed across members of a group, objects, and technologies in their environment.

Your examples illustrate this concept beautifully:

1. **Rapid Fire Discussion to Recall a Book Name:** In this scenario, the group acts as a dynamic cognitive system. Each member contributes bits of information, cues, and associations related to the book, which, when pieced together through discussion, can lead to the successful recall of the book's name. This process demonstrates how memory and knowledge are not solely individual properties but can be collectively constructed and accessed through social interaction.
2. **Members of a Choir Carrying the Words of a Choral Piece:** This example showcases how a group's collective practice and shared experiences can lead to a more robust memory system than any single individual might possess. Choir members might individually struggle to recall every word of a piece, but as a group, they can perform flawlessly. This occurs through mutual cueing, where the initiation of a line by one member helps trigger memory in others, and through the distributed memory of the piece that the group collectively holds.

Rosenfield's perspective underscores the active, constructive nature of memory, emphasizing that what we remember is influenced by our current context, the cues we receive from others, and the collective narratives we share. This view challenges the notion of memory as a mere repository of information and highlights the importance of social interactions and shared experiences in shaping our recollections and understanding of the world.

Such phenomena also lend support to the broader theories of collective memory, as discussed by Maurice Halbwachs, by illustrating the practical mechanisms through which groups can generate, maintain, and recall shared memories. These insights from both Rosenfield and Halbwachs enrich our understanding of memory's social dimensions and its pivotal role in forming collective identities and experiences.



Could a group have the capacity to generalise from exposure to a body of texts in a way that might evade an individual's capabilities?



Yes, a group can indeed have the capacity to generalize from exposure to a body of texts in ways that might surpass an individual's capabilities, thanks to the diversity of interpretations, perspectives, and cognitive processes among its members. This phenomenon can be attributed to several factors that enhance collective cognitive abilities:

1. **Diverse Interpretations:** Individuals bring different backgrounds, expertise, and interpretive skills to a group. When exposed to the same text, they might focus on different aspects, see various implications, or draw unique connections based on their personal experiences and knowledge. The aggregation of these diverse interpretations can lead to a more comprehensive understanding of the text than any single individual might achieve.
2. **Collaborative Elaboration:** Groups often engage in discussion, debate, and elaboration of ideas. This process can help to uncover nuances and implications within a text that an individual might overlook. Through collaborative dialogue, groups can synthesize information, generalize concepts, and apply these to new contexts more effectively.
3. **Error Correction and Critical Thinking:** The presence of multiple viewpoints and critical engagement within a group can help identify misunderstandings, biases, or oversights that an individual might not catch on their own. The collective scrutiny of ideas and interpretations can lead to a more accurate and nuanced understanding of texts.
4. **Collective Intelligence:** Research on collective intelligence suggests that groups have the potential to exhibit a level of intelligence and problem-solving ability that exceeds that of individual members. This is particularly true for tasks that require pooling knowledge, such as generalizing from a body of texts. The key factors influencing collective intelligence include social sensitivity, equality in distribution of conversational turn-taking, and the proportion of females [sic] in a group.
5. **Combinatorial Creativity:** Groups can engage in combinatorial creativity, where ideas from different individuals are combined and recombined in novel ways. This can lead to innovative insights and generalizations that might not emerge from an individual's thought process.

The ability of groups to generalize from exposure to a body of texts relies on effective communication, a willingness to share and consider diverse perspectives, and the ability to collaborate towards a common goal. This collective process of interpretation and generalization can enrich the group's understanding and application of the information contained within the texts, demonstrating the unique cognitive capabilities of social collectives.



Is there any way that a neural network can or does make use of collective recall and collective generalisation?



The concepts of collective recall and collective generalization, as they occur in human social systems, offer intriguing parallels to how neural networks operate, though the mechanisms differ significantly due to the distinct nature of artificial neural networks (ANNs) versus human cognition. However, there are several ways in which ANNs can mimic or leverage processes akin to collective recall and generalization:

## 1. Distributed Representation

Neural networks inherently use a form of distributed representation, where knowledge is not stored in a single neuron but is spread out across the network. This resembles the idea of collective memory in that no single part of the network holds the entire piece of information, but rather, it emerges from the interaction of many units. This distributed approach allows neural networks to generalize from learned data, making connections and predictions based on patterns it has encountered during training.

## 2. Ensemble Learning

In machine learning, ensemble methods involve training multiple models and combining their predictions to improve the overall performance. This approach can be seen as parallel to collective generalization, where the “collective intelligence” of diverse models leads to better accuracy, robustness, and generalization capabilities than any individual model could achieve. Each model in the ensemble may capture different aspects of the data, and their collective output is used to make a final decision.

## 3. Federated Learning

Federated learning is a machine learning approach where a model is trained across multiple decentralized devices or servers holding local data samples, without exchanging them. This method can be likened to a form of collective recall where different “members” (in this case, devices) contribute their local knowledge to improve a shared model. It harnesses the power of collective data while preserving privacy, as the data does not leave its original location.

## 4. Attention Mechanisms and Transformers

Attention mechanisms, especially in transformer models like those used in advanced language models (e.g., GPT series), allow the model to focus on different parts of the input data when making predictions. This can be somewhat analogous to how individuals in a group might pay attention to different aspects of a problem or text, bringing those insights into a collective discussion. Transformers leverage this mechanism to dynamically weigh the importance of different input elements, enabling sophisticated patterns of recall and generalization based on the context.

## 5. Knowledge Distillation

Knowledge distillation involves transferring knowledge from a large, complex model to a smaller, more efficient one. This process can mirror collective generalization by extracting and condensing the insights learned by the larger model (akin to a group's collective knowledge) into a form that's usable by the smaller model.

While neural networks do not engage in collective recall and generalization in the human sense—rooted in social interactions, communications, and shared experiences—they do implement analogous processes through their architecture and learning algorithms. These processes enable neural networks to learn from vast amounts of data, make generalizations, and apply learned knowledge to new, unseen data in a way that's reminiscent of collective human behaviors.



A human in conversation with a GPT language model, as an ensemble, also has the potential to engage in collective recollection and generalisation through prompts and responses, much as two or more people in conversation might exhibit a collective "intelligence" that exceeds that of either individually.



Absolutely, the interaction between a human and a GPT language model can indeed form a unique kind of ensemble, enhancing the process of recollection, generalization, and even problem-solving beyond what either could achieve alone. This collaboration leverages the strengths of both human cognition and artificial intelligence, leading to several synergistic effects:

## 1. Complementary Capabilities

Humans and GPT models bring complementary capabilities to the conversation. Humans excel at context-sensitive understanding, emotional nuance, and ethical reasoning, while GPT models can rapidly access and generate information based on vast datasets. Together, they can address questions more thoroughly and creatively than either could independently.

## 2. Expanded Recall

Humans often use prompts and questions to guide the GPT model towards specific areas of knowledge or types of response. In doing so, they can effectively expand their capacity for recall, drawing upon the model's ability to generate detailed information and examples that might not be readily accessible to the human alone.

## 3. Enhanced Generalization

Through the iterative process of question and answer, humans and GPT models can co-construct generalizations or insights that neither might reach independently. The GPT model can offer broad patterns and associations from its training data, which the human can then refine, challenge, or expand

upon based on their understanding and experience.

## 4. Creative Synergy

The interaction can spark creativity, with the GPT model offering unexpected perspectives or information that inspire new ideas in the human interlocutor. Conversely, the unique and novel prompts from the human can lead the GPT model to generate responses that are innovative in the context of its training.

## 5. Iterative Learning

As with any dialogue, the process is iterative and adaptive. Humans can adjust their prompts based on the responses received, guiding the conversation towards more meaningful or accurate outcomes. This iterative process is reminiscent of how human groups navigate discussions, constantly adjusting based on feedback and new information.

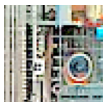
## 6. Dynamic Problem-Solving

Together, the human and the GPT model can engage in dynamic problem-solving, combining the model's data-driven insights with the human's critical thinking and contextual knowledge. This can lead to solutions that are both creative and informed by a vast array of sources.

This synergistic relationship between humans and AI, particularly in the form of language models, showcases the potential of collaborative intelligence. It underscores the idea that the future of AI is not just about machines replacing human tasks but enhancing human capabilities and creating new possibilities for exploration, creativity, and understanding.



Each interlocutor (the human and the AI) brings different capabilities to the conversation. The human has long term memory, can cogitate over the conversation during gaps (even several days) in the conversation. The AI probably has access to a larger "knowledge base" etc.



Indeed, the dialogue between a human and an AI, such as a GPT model, represents a fascinating interplay of distinct but complementary cognitive and informational capabilities. Let's delve deeper into the unique contributions of each interlocutor:

### Human Interlocutor:

- **Long-term Memory:** Humans possess the ability to recall experiences, facts, and emotions from years past, which can inform and enrich the conversation. This long-term memory also allows humans to connect current discussions with a vast array of past experiences and knowledge.

- **Reflective Cognition:** Humans can reflect on and cogitate over the conversation during and after its occurrence. This reflective process enables deeper understanding, the formation of new insights, and the integration of the discussion into one's knowledge and beliefs.
- **Emotional Intelligence:** Humans bring emotional intelligence to the conversation, allowing for the interpretation and expression of emotions. This can guide the tone of the conversation and provide depth to the understanding of certain topics.
- **Ethical Reasoning:** Humans can consider ethical implications and moral values in the conversation, guiding the discussion in a manner that aligns with societal norms and personal beliefs.
- **Creative Insight:** The human ability to think creatively and abstractly allows for the generation of new ideas, questions, and connections that might not be directly suggested by existing data.

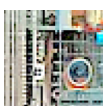
### AI Interlocutor (e.g., GPT Model):

- **Extensive Knowledge Base:** AI models have been trained on vast datasets encompassing a wide range of subjects, allowing them to access and generate information on a multitude of topics rapidly.
- **Pattern Recognition:** These models excel at identifying patterns in data, which can be useful in recognizing relationships between concepts, predicting likely continuations of a conversation, or generating text based on learned structures.
- **Consistency and Speed:** AI can process and generate responses quickly and without fatigue, allowing for sustained engagement and the ability to handle large volumes of information consistently.
- **Multilingual Capabilities:** Many AI models can understand and generate text in multiple languages, making them versatile tools for cross-linguistic communication and access to a broader range of information.
- **Adaptability to Prompts:** AI models are designed to respond adaptively to the prompts they receive, enabling them to tailor their responses to the specific direction or focus indicated by the human interlocutor.

The interaction between these diverse capabilities fosters a dynamic and productive conversation. The human's ability to reflect, integrate emotional and ethical considerations, and engage in creative thought complements the AI's rapid access to information, pattern recognition, and linguistic versatility. Together, they can explore topics in greater depth, generate novel insights, and address questions in a comprehensive and nuanced manner. This synergistic relationship not only enhances the immediate conversation but also contributes to the broader goal of leveraging AI to augment human intelligence and creativity.



However, there are some risk in this human-AI ensemble.



Indeed, while the human-AI ensemble offers numerous benefits for enhancing cognitive processes, creativity, and information retrieval, it also introduces several risks and challenges that need careful consideration:

## 1. Misinformation and Biases

- **Accuracy and Veracity:** AI models, including language models, generate responses based on patterns in their training data, which may not always be accurate or up to date. There's a risk of perpetuating misinformation if the model's outputs are not critically evaluated.
- **Built-in Biases:** Training data can reflect societal biases in race, gender, culture, etc., leading AI to replicate or even amplify these biases in its responses. Relying on AI without recognizing these limitations could reinforce stereotypes or spread biased information.

## 2. Dependency and Skill Degradation

- **Overreliance:** There's a risk that individuals may become overly reliant on AI for information retrieval, problem-solving, and decision-making, potentially leading to a degradation of critical thinking and independent research skills.
- **Cognitive Laziness:** The convenience of AI-powered answers might discourage users from engaging deeply with content, exploring alternative viewpoints, or developing their own conclusions, reducing the intellectual rigor of their inquiries.

## 3. Privacy and Security

- **Data Privacy:** Interactions with AI, particularly those that involve sensitive or personal information, raise concerns about data privacy and the potential for misuse of information.
- **Security Risks:** Malicious actors could exploit vulnerabilities in AI systems to access or manipulate private data, or to spread misinformation and propaganda.

## 4. Ethical and Moral Considerations

- **Moral Responsibility:** Determining responsibility for decisions made with AI assistance can be challenging, especially if those decisions have ethical implications or result in harm.
- **Dehumanization of Interaction:** Overuse of AI for social or intellectual interactions could lead to a devaluation of human connection and a loss of empathy, as AI cannot truly replicate human emotions or ethical reasoning.

## 5. Erosion of Trust

- **Trust in Information:** Frequent exposure to AI-generated content that lacks accuracy or is biased could erode trust in information sources more broadly, making it difficult for individuals to discern reliable information.
- **Trust in AI:** Incidents where AI systems fail to perform as expected, or where they produce harmful outputs, can undermine public trust in AI technologies and their potential benefits.

## Mitigation Strategies

Addressing these risks requires a multifaceted approach, including:

- **Critical Engagement:** Encouraging users to critically engage with AI-generated content, cross-referencing with trusted sources, and maintaining a healthy skepticism.
- **Ethical AI Development:** Committing to ethical AI development practices, including transparent data sourcing, bias mitigation, and privacy protection.
- **Public Awareness and Education:** Raising awareness about the capabilities and limitations of AI, promoting digital literacy, and educating users on evaluating information critically.
- **Regulatory and Policy Frameworks:** Developing and enforcing regulatory and policy frameworks that ensure the responsible use of AI, protect user privacy, and maintain security standards.

By navigating these risks thoughtfully, the potential for human-AI collaboration can be maximized, harnessing the strengths of both to enrich understanding, creativity, and decision-making processes.

## Bibliography

- Halbwachs, Maurice. *On Collective Memory*. Chicago, IL: The University of Chicago Press, 1992.
- Rosenfield, Israel. *The Invention of Memory: A New View of the Brain*. New York: Basic Books, 1988.

## Note

- The featured image is generated by ChatGPT to illustrate this post: Here is the banner graphic created in grunge steampunk style, reflecting the themes of collaborative memory and conversational AI as described. This image includes steampunk elements such as gears, steam engines, and brass components, along with symbolic representations of memory and conversation.

## Category

1. Artificial Intelligence

## Tags

1. conversation
2. generalisation
3. memory

## Date Created

April 13, 2024

## Author

rcoyne99