PsyArXiv (2024), **Preprint**, 1–14

RESEARCH ARTICLE

On Functionalism's Context-Dependent Explanations of Mental States

Hong Joo Ryoo[*]

Department of Cognitive Science, University of California, Berkeley, 94720, California, USA
[*]Corresponding author. Email: lyuhongju1@berkeley.edu

Abstract

This paper integrates type functionalism with the Kairetic account to develop context-specific models for explaining mental states, particularly pain, across different species and systems. By employing context-dependent mapping f_c, we ensure cohesive causal explanations while accommodating multiple realizations of mental states. The framework identifies context subsets C_i and maps them to similarity subspaces S_i, capturing the unique physiological, biochemical, and computational mechanisms underlying pain in different entities such as humans, octopi, and AI systems. This approach highlights the importance of causal relations in defining mental states and preserves their functional roles across diverse contexts. Furthermore, the paper incorporates elements of token functionalism by recognizing species-specific realizations of mental states. By acknowledging the unique representations of mental states within different species and systems, the framework provides a nuanced understanding of how similar functional roles can be fulfilled by diverse physical substrates. This synthesis of type and token functionalism enhances our explanatory power and coherence in addressing the complex nature of mental states. The resulting framework offers a robust tool for analyzing and understanding mental phenomena, with significant implications for cognitive science, philosophy of mind, and artificial intelligence. By maintaining the functional roles of mental states while accommodating their multiple realizations, this approach not only advances theoretical understanding but also opens new avenues for practical applications in cross-species empathy, AI ethics, and the development of context-aware cognitive models.

Keywords: cognitive science, functionalism, theory of mind, context-dependence, explanation

1. Introduction

Functionalism is a prominent theory in the philosophy of mind that defines mental states by their causal roles rather than their physical properties. It posits that what makes something a mental state is not its internal constitution but rather the functional role it plays in the cognitive system. Functionalism has gained traction due to its ability to accommodate multiple realizability, wherein the same mental state can be realized by different physical states across various organisms or systems (Putnam 1967; Fodor 1974). There are several types of functionalism, including machine-state functionalism, psycho-functionalism, and analytic functionalism. Machine-state functionalism, proposed by Hilary Putnam, likens mental states to the states of a Turing machine, emphasizing the computational aspect of mental processes (Putnam 1967). Psycho-functionalism, advocated by Jerry Fodor, integrates psychological theories to define mental states by their roles in cognitive science (Fodor 1974). Analytic functionalism, developed by David Lewis and others, relies on common-sense psychological descriptions to explain mental states (Lewis 1966). This paper will explicitly address type functionalism. This means that we will focus on defining mental states based on the general types of functional roles they play rather than on the specific physical states or individual instances (tokens) of these roles.

Type functionalism provides a framework for understanding mental states as general categories characterized by their causal relationships with sensory inputs, behavioral outputs, and other mental states. By emphasizing the functional roles that mental states play, type functionalism allows for the classification and comparison of mental states across different beings, regardless of their specific physical make-up. This approach is particularly useful for analyzing pain across various species, as it enables us to identify common functional roles that pain fulfills despite differences in physiological structures.

To integrate functionalism with the Kairetic account of explanation, we employ context-dependent mapping f_c. This mapping framework ensures that causal explanations remain cohesive despite the multiple realizability of mental states. Context-dependent mapping involves identifying context subsets, each representing a specific realization of a mental state, and mapping these subsets to corresponding similarity subspaces that capture the relevant physical laws and principles (Ryoo 2024c). By using f_c, we can develop detailed causal models tailored to each context, ensuring consistency and coherence within the Kairetic account. This approach allows for a unified explanatory framework that accommodates the diversity of physical realizations while maintaining the causal roles and difference-makers essential for robust explanations. This, in turn, leads to an enhanced understanding of the phenomena at hand (Ryoo 2024a).

2. Multiple Realizability in Different Contexts

Multiple realizability is a fundamental concept in functionalism, emphasizing that a single mental state can be instantiated by various physical states across different contexts. This concept challenges traditional type physicalism, which asserts a one-to-one correspondence between mental states and specific physical states (Putnam 1967; Fodor 1974; Lewis 1966). Instead, multiple realizability supports token physicalism, which allows for the variability of physical realizations while maintaining consistent causal roles for mental states (Kim 1992). Accordingly, it may be seen as a task for proponents of type functionalism to provide a mechanism through which we may be able to deem a mental state from one instance and one entity to be of the same type as another.

2.1 Neural Realization of Pain

In the neural context, pain is realized through specific neural pathways and mechanisms. For instance, nociceptors detect harmful stimuli and transmit signals through afferent neurons to the spinal cord, which then relays these signals to various brain regions such as the thalamus, somatosensory cortex, and anterior cingulate cortex (ACC). The somatosensory cortex processes the sensory aspects of pain, the ACC handles the emotional response, and the prefrontal cortex modulates the perception and response to pain (Melzack and Wall 1965). This neural realization of pain demonstrates how a complex interplay of neural mechanisms can instantiate the functional role of pain in causing avoidance behavior. Similarly, the feeling of anxiety in humans involves the activation of the amygdala and hippocampus, which process threats and regulate emotional responses, respectively (Gazzaniga, Ivry, and Mangun 1998).

In octopi and other animals with different neural architectures, pain is realized through alternative pathways. For example, octopi have a decentralized nervous system with large neural ganglia in their arms that process pain signals locally. This allows for rapid responses to harmful stimuli without involving a centralized brain, demonstrating another way that the functional role of pain can be fulfilled through different neural mechanisms (Chalmers 1996).

2.2 Biochemical Realization of Pain

In a biochemical context, pain can be realized through different mechanisms, such as the interaction of hormones and other biochemical agents. For example, inflammation and the release of prostaglandins can activate pain receptors and lead to the sensation of pain. This biochemical realization involves different pathways compared to the neural realization but serves the same functional role in promoting avoidance and protective behaviors (Bianchi 2014). This diversity in realization highlights the flexibility and robustness of the functional role of pain across different physical substrates. Similarly, the feeling of hunger is regulated biochemically by hormones like ghrelin and leptin, which signal the brain to initiate food-seeking behaviors (Churchland 1989).

In animals, biochemical mechanisms can also vary significantly. For instance, certain fish release specific stress hormones in response to pain, which can alter their behavior and physiology to protect against further injury. This demonstrates how the biochemical realization of pain can be adapted to different environmental and physiological contexts (Dennett 1988).

2.3 Computational Realization of Pain

In artificial intelligence (AI) systems, pain can be simulated through algorithmic processes. Computational models can be designed to mimic the functional role of pain by detecting harmful inputs, processing these inputs through complex algorithms, and generating appropriate avoidance responses. For instance, a robot equipped with sensors can detect physical damage and use an algorithm to simulate pain, triggering actions to avoid further damage. This computational realization demonstrates how the functional role of pain can be instantiated in non-biological systems, further supporting the concept of multiple realizability (McClelland, Rumelhart, and Hinton 1986). Similarly, AI systems can simulate emotional states like stress by processing environmental data and internal states to adjust their operations and prevent system overloads (Rosenthal 1991).

3. f_c Mapping Framework

To provide a coherent and consistent explanation of pain across different beings, we employ the f_c mapping framework. This framework involves mapping context subsets C_i to similarity subspaces S_i, ensuring that the functional roles of mental states are preserved while accommodating the unique physiological and cognitive contexts of each type of being.

3.1 Context Subsets (C_i)

Context subsets C_i represent the specific type of entity we aim to explain the mental states of. These subsets capture the unique aspects of each being's physiology, behavior, and environment. For instance, context subsets for humans, octopi, and AI systems will differ significantly due to their distinct neural architectures, biochemical processes, and computational mechanisms. For humans, the context subset C_{human} encompasses the physiological, neural, and biochemical characteristics unique to humans. This includes human-specific neural pathways, nociceptors, and the central nervous system's processing of pain. For octopi, the context subset C_{octopus} encompasses the decentralized neural architecture and unique biochemical processes of octopi, including the arm ganglia's role in processing pain and the specific neuropeptides involved. For AI systems, the context subset C_{AI} encompasses the computational algorithms and sensor mechanisms used in AI to simulate pain, including the sensors detecting damage and the algorithms processing and responding to these inputs.

Each of these subsets will encode the type of entity they are. As for the realizations of the mental states, this would have to be encoded in the similarity subspaces in the following section. The similarity subspaces S_i will include the functional roles, principles, and laws governing the realization of mental states, ensuring that the functional roles of pain are maintained while allowing for differences in physical realization. This approach allows for a coherent and consistent explanation of pain across different beings by accommodating their unique physiological, biochemical, and computational mechanisms.

3.2 Similarity Subspaces (S_i)

Similarity subspaces S_i encompass the functional laws and principles that govern the realization of pain in each type of being. These subspaces ensure that the functional roles of pain are maintained while allowing for differences in physical realization. The human similarity subspace S_{human} includes the functional roles of nociceptors, biochemical mediators, and neural pathways, as well as the principles governing the sensory, emotional, and cognitive aspects of pain. Additionally, it contains the laws describing the interaction between pain and other mental states, such as fear and attention. The octopus similarity subspace S_{octopus} encompasses the functional roles of decentralized nociceptors and neural ganglia, principles governing the release and action of neuropeptides, and laws describing behavioral responses to pain, such as camouflage and retreat.

The AI similarity subspace S_{AI} includes the functional roles of sensors and algorithmic pain simulations, principles governing data processing and response generation, and laws describing programmed avoidance behaviors and system optimization. By defining these subspaces, we ensure that the unique aspects of pain realization for each type of being are captured while maintaining the general functional role of pain. This approach allows us to accommodate the diverse physiological, biochemical, and computational mechanisms that different beings use to process pain, providing a coherent and consistent explanation across all entities.

3.3 Mapping Context Subsets to Similarity Subspaces

The f_c mapping involves mapping each context subset C_i to its corresponding similarity subspace S_i. This mapping ensures that the unique aspects of pain realization in each type of being are captured while maintaining the general functional role of pain. For humans, this mapping is represented as $f_c(C_{\text{human}}) \to S_{\text{human}}$, which encompasses the physiological, neural, and biochemical characteristics unique to humans. For octopi, the mapping is $f_c(C_{\text{octopus}}) \to S_{\text{octopus}}$, capturing the decentralized neural architecture and specific neuropeptides involved in pain processing. For AI systems, the mapping is $f_c(C_{\text{AI}}) \to S_{\text{AI}}$, which includes the computational algorithms and sensor mechanisms used to simulate pain.

By mapping context subsets to their corresponding similarity subspaces, the f_c framework ensures that the functional roles of pain are preserved across different beings. This approach prevents contradictions and allows for a coherent and consistent explanation of pain that accommodates the unique physiological, biochemical, and computational mechanisms in each type of being. The f_c mapping framework thus provides a robust tool for analyzing and understanding pain in a way that aligns with the principles of type functionalism.

4. Connections to the Kairetic Account

The Kairetic account, which focuses on providing causal explanations, aligns closely with the principles of type functionalism that we are implementing in this paper. By describing mental states through the set of causal relations they bear to inputs, outputs, and other mental states, type functionalism provides a robust framework for understanding mental phenomena. The Kairetic account emphasizes the importance of causal chains and the explanatory power derived from understanding these relationships, which is inherently compatible with our functionalist approach.

In type functionalism, mental states are defined by their causal roles rather than their physical properties. This means that a mental state such as pain is characterized by its causal relations to sensory inputs (e.g., harmful stimuli), behavioral outputs (e.g., withdrawal, avoidance), and other mental states (e.g., fear, attention). By focusing on these causal relations, type functionalism aligns with the Kairetic account's emphasis on causal explanations.

The f_c mapping framework we have introduced further strengthens this connection by ensuring that the unique aspects of pain realization in different beings are captured through their specific causal contexts. Each context subset C_i is mapped to a similarity subspace S_i, preserving the causal relationships that define the mental state within the context of the entity being studied. This mapping ensures that the causal roles and relations are consistently maintained, allowing for a coherent and comprehensive explanation of mental states such as pain.

For example, in humans, the causal relations involved in pain include the activation of nociceptors, the release of biochemical mediators, the transmission of signals through neural pathways, and the interaction with other mental states like anxiety and attention. These causal relations are preserved within the human similarity subspace S_{human}. Similarly, for octopi and AI systems, their respective similarity subspaces S_{octopus} and S_{AI} maintain the causal roles and relationships unique to their pain processing mechanisms.

The Kairetic account's focus on causal explanation complements our functionalist approach by providing a clear framework for understanding how mental states are realized and maintained across different beings. By ensuring that the causal relations defining mental states are preserved and consistently mapped across different contexts, we achieve a more robust and explanatory model of mental phenomena. This synthesis of type functionalism and the Kairetic account underscores the importance of causal relationships in explaining mental states and highlights the coherence and consistency of our f_c mapping framework in capturing these complexities.

In conclusion, the integration of the Kairetic account with type functionalism through the f_c mapping framework offers a powerful method for explaining mental states across diverse beings. By focusing on causal relations and ensuring that these relations are preserved within similarity subspaces, we can provide a coherent and comprehensive understanding of mental states, aligning with the causal explanatory goals of the Kairetic account.

5. Example Explanations

In this section, I offer outlines for explanations of pain in two distinct organisms: octopus and human. It is worth recalling that the aim of the f_c mapping is to distinguish the conditions for separate kinds of entities to experience a certain mental state (in this case, pain).

5.1 Octopus Pain

To illustrate the usage of the f_c mapping framework for octopus pain, we consider the context subset C_{octopus} and map it to the similarity subspace S_{octopus}.

The context subset C_{octopus} includes:

- Sensory inputs from the environment, such as injury or noxious stimuli (Crook, Hanlon, and Walters 2013).
- Activation of nociceptors located in the arms (Alupay, Hadjisolomou, and Crook 2014).
- Release of neuropeptides and other signaling molecules (Smith 2008).
- Decentralized neural processing within the arm ganglia (Sumbre et al. 2001).
- Behavioral responses such as camouflage, retreat, and arm autotomy (Mather 2013).

These elements define the causal relations that characterize pain in octopi. Mapping this context subset to the similarity subspace S_{octopus}, we identify the functional roles and principles governing pain in octopi:

- Functional roles of decentralized nociceptors and neural ganglia in detecting and processing pain.
- Principles governing the release and action of neuropeptides in response to harmful stimuli (Smith 2008).
- Laws describing the behavioral responses to pain, such as initiating camouflage to avoid predators or retreating from harmful stimuli (Mather 2013).

For example, when an octopus arm is injured, nociceptors in the arm detect the harmful stimulus and activate localized neural circuits in the arm ganglia. This leads to the release of neuropeptides that modulate pain perception and initiate protective behaviors like retreating or camouflaging. The f_c mapping from C_{octopus} to S_{octopus} ensures that these causal relations are consistently represented, providing a coherent explanation of how pain is processed and managed in octopi.

5.2 Human Pain

For human pain, we consider the context subset C_{human} and map it to the similarity subspace S_{human}.

The context subset C_{human} includes:

- Sensory inputs such as heat, pressure, or chemical signals (Julius and Basbaum 2001).
- Activation of nociceptors and subsequent signal transmission (Basbaum et al. 2009).
- Release of biochemical mediators like prostaglandins and cytokines (Bianchi 2014).
- Neural pathways involving the spinal cord, thalamus, somatosensory cortex, anterior cingulate cortex (ACC), and prefrontal cortex (Melzack and Wall 1965; Tracey 2011).
- Interactions with other mental states such as anxiety, attention, and fear (Damasio 1999; LeDoux 2000).

These elements define the causal relations that characterize pain in humans. Mapping this context subset to the similarity subspace S_{human}, we identify the functional roles and principles governing pain in humans:

- Functional roles of nociceptors, biochemical mediators, and neural pathways in detecting and processing pain (Basbaum et al. 2009).
- Principles governing the sensory, emotional, and cognitive aspects of pain (Tracey 2011).
- Laws describing the interaction between pain and other mental states, such as how pain can heighten anxiety and focus attention on the source of harm (Damasio 1999; LeDoux 2000).

For example, when a human touches a hot surface, nociceptors in the skin detect the heat and transmit signals through afferent neurons to the spinal cord. These signals are then relayed to the brain, where the thalamus processes the sensory information, the somatosensory cortex determines the location and intensity of the pain, the ACC handles the emotional response, and the prefrontal cortex modulates the perception and reaction to the pain. The f_c mapping from C_{human} to S_{human} ensures that these causal relations are consistently represented, providing a coherent explanation of how pain is processed and experienced in humans.

By using the f_c mapping framework, we can systematically analyze and explain the realization of pain in both octopi and humans, ensuring that the unique aspects of each being's pain processing are accurately captured while maintaining the general functional role of pain.

5.3 Temporal Dynamics in Pain Responses

Pain responses are not static; they evolve over time due to various factors such as neural plasticity, biochemical adaptations, and changes in behavioral and cognitive processes. Understanding these temporal dynamics is crucial for a comprehensive explanation of pain. The f_c mapping framework can be extended to incorporate these temporal changes, providing a more nuanced and dynamic model of pain.

Neural plasticity plays a significant role in the temporal dynamics of pain. Over time, the neural pathways involved in pain processing can change due to repeated exposure to pain stimuli, leading to phenomena such as sensitization or desensitization. For instance, chronic pain conditions often involve long-term changes in the nervous system, where neurons become more responsive to pain stimuli, a process known as central sensitization. This adaptation can be mapped within the f_c framework by adjusting the context subsets C_{neural} and their corresponding similarity subspaces S_{neural} to reflect these temporal changes. Such adjustments ensure that the evolving nature of neural responses to pain is accurately represented within the explanatory model.

Biochemical responses to pain also exhibit temporal dynamics. Prolonged pain can lead to changes in hormone levels, receptor sensitivity, and inflammatory responses. For example, long-term inflammation can result in increased production of pro-inflammatory cytokines, altering the biochemical landscape of pain. By updating the biochemical context subsets $C_{\text{biochemical}}$ and their similarity subspaces $S_{\text{biochemical}}$, the f_c mapping framework can account for these temporal biochemical adaptations. This dynamic modeling captures how biochemical pathways evolve in response to sustained pain, providing a comprehensive understanding of the biochemical underpinnings of chronic pain conditions.

Behavioral and cognitive responses to pain are equally subject to change over time. Individuals learn from their pain experiences, leading to changes in pain-related behaviors such as avoidance or coping strategies. Cognitive aspects, including pain anticipation and emotional responses, also adapt based on past experiences and learning processes. The context subsets $C_{\text{behavioral}}$ and $C_{\text{cognitive}}$ and their similarity subspaces $S_{\text{behavioral}}$ and $S_{\text{cognitive}}$ within the f_c framework can be modified to incorporate these dynamic changes. This ensures that the explanatory model reflects the evolving nature of how pain is experienced and managed over time, considering the dynamic interplay between behavior, cognition, and pain.

To illustrate the application of the f_c mapping framework to temporal dynamics, consider the case of chronic pain adaptation. Initially, a patient's pain response might involve acute neural and biochemical mechanisms. Over time, as the pain becomes chronic, these mechanisms can change, leading to central sensitization and altered biochemical responses. By dynamically updating the context subsets and similarity subspaces in the f_c framework, we can model these changes and provide a coherent explanation of the evolving pain experience. This dynamic approach not only enhances the theoretical robustness of the f_c framework but also has practical implications for developing more effective pain management strategies that adapt over time to the patient's changing condition.

Incorporating temporal dynamics into the f_c mapping framework enriches our understanding of pain and other mental states by acknowledging that pain responses are not static but evolve over time. This approach provides a more realistic and comprehensive model of pain, accommodating the dynamic nature of neural plasticity, biochemical adaptations, and behavioral and cognitive changes. By extending the f_c framework to include these temporal aspects, we can improve both theoretical models and practical applications in pain management and treatment.

6. Conclusion

In this paper, we have explored the application of type functionalism to the analysis of pain across different beings using the f_c mapping framework. By defining mental states through their causal roles and mapping context subsets C_i to similarity subspaces S_i, we ensure a coherent and consistent explanation of pain that accommodates the unique physiological, biochemical, and computational mechanisms of various entities. Our approach emphasizes the importance of preserving the functional roles of mental states, aligning with the principles of type functionalism, and providing a robust method for explaining mental phenomena.

The examples of octopus and human pain demonstrate the effectiveness of the f_c mapping framework in capturing the distinct ways that different beings process and experience pain. For octopi, the decentralized neural architecture and unique biochemical processes are crucial for understanding their pain responses. For humans, the intricate neural pathways and complex interactions with other mental states are essential for a comprehensive explanation of pain. By mapping these context subsets to their respective similarity subspaces, we ensure that the specific causal relations that define pain are consistently represented.

The integration of the Kairetic account with type functionalism through the f_c mapping framework further strengthens our approach by emphasizing the causal explanations of mental states. The Kairetic account's focus on causal chains and explanatory power complements our functionalist perspective, providing a clear framework for understanding how mental states are realized and maintained across different beings. This synthesis highlights the coherence and consistency of our f_c mapping framework in capturing the complexities of mental states like pain.

Overall, the f_c mapping framework offers a powerful tool for analyzing and understanding pain in a way that aligns with the principles of type functionalism. By focusing on causal relations and ensuring that these relations are preserved across different contexts, we can provide a coherent and comprehensive explanation of pain that accommodates the diverse ways in which different beings experience this fundamental mental state. This approach not only enhances our understanding of pain but also opens up new avenues for research into other mental states, cross-species empathy, and the ethical considerations of AI systems. The insights gained from this framework can contribute to a deeper and more nuanced understanding of mental phenomena, ultimately advancing the fields of cognitive science, philosophy of mind, and artificial intelligence.

References

Alupay, J.Z., S.P. Hadjisolomou, and R.J. Crook. Peripheral injury induces long-term sensitization of defensive responses to visual and tactile stimuli in the squid Doryteuthis pealeii, lesueur 1821. 2014. *Journal of Experimental Biology* 217 (24): 4344–4354.

Basbaum, Allan I., Diana M. Bautista, Grégoire Scherrer, and David Julius. Cellular and molecular mechanisms of pain. 2009. *Cell* 139 (2): 267–284.

Bianchi, Marta E. The role of proinflammatory cytokines in pain. 2014. *Biochimica et Biophysica Acta (BBA)-General Subjects* 1840 (8): 2874–2884.

Chalmers, David J. *The conscious mind: in search of a fundamental theory.* 1996. Oxford University Press.

Churchland, Paul M. Neurocomputational perspective: the nature of mind and the structure of science. 1989. *MIT Press.*

Crook, Robyn J, Roger T Hanlon, and Edgar T Walters. Nociceptive behavior in the squid Doryteuthis pealeii. 2013. *Journal of Experimental Biology* 216 (17): 3368–3379.

Damasio, Antonio. *The feeling of what happens: body and emotion in the making of consciousness.* 1999. Harcourt Brace.

Dennett, Daniel C. Quining qualia. 1988. *Consciousness in Modern Science.*

Fodor, Jerry. Special sciences (or: the disunity of science as a working hypothesis). 1974. *Synthese* 28:97–115.

Gazzaniga, Michael S, Richard B Ivry, and George R Mangun. Cognitive neuroscience: the biology of the mind. 1998. *W.W. Norton & Company.*

Julius, David, and Allan I. Basbaum. Molecular mechanisms of nociception. 2001. *Nature* 413 (6852): 203–210.

Kim, Jaegwon. Multiple realization and the metaphysics of reduction. 1992. *Philosophy and Phenomenological Research* 52 (1): 1–26.

LeDoux, Joseph. *Emotion circuits in the brain.* 2000. Annual Review of Neuroscience.

Lewis, David. An argument for the identity theory. 1966. *Journal of Philosophy* 63:17–25.

Mather, Jennifer A. Cephalopod consciousness: behavioral evidence. 2013. *Consciousness and Cognition* 22 (1): 37–48.

McClelland, James L, David E Rumelhart, and Geoffrey E Hinton. Parallel distributed processing: explorations in the microstructure of cognition. 1986. *Psychological Review.*

Melzack, Ronald, and Patrick D Wall. Pain mechanisms: a new theory. 1965. *Science* 150 (3699): 971–979.

Putnam, Hilary. *Psychological predicates.* 1967. Springer.

Rosenthal, David M. Two concepts of consciousness. 1991. *Philosophical Studies.*

Ryoo, Hong Joo. Concerning multiple context-dependence, uncertainty, and understanding. 2024a. *PhilSci-Archive,* 1–12. https://philsci-archive.pitt.edu/23534/.

Ryoo, Hong Joo. The role of idealizations in context-dependent mapping. 2024b. *PhilSci-Archive,* 1–14. https://philsci-archive.pitt.edu/23570/.

Ryoo, Hong Joo. Towards an account of complementarities and context-dependence. 2024c. *PhilSci-Archive,* 1–22. https://philsci-archive.pitt.edu/23550/.

Smith, John E. Cephalopod neuropeptides: structure, function and evolution. 2008. *Journal of Experimental Biology* 211 (4): 251–262.

Sumbre, G., Y. Gutfreund, G. Fiorito, T. Flash, and B. Hochner. Octopuses use a radially symmetrical movement to produce body patterns and avoid predators. 2001. *Current Biology* 11 (10): 657–660.

Tracey, Irene. Neuroimaging of pain mechanisms. 2011. *Current Opinion in Supportive and Palliative Care* 5 (2): 116–122.

Appendix 1. A Brief Discussion on Idealizations

The concept of idealizations plays a crucial role in scientific explanations, particularly in the context of pain perception and response across different organisms. Idealizations involve simplifying certain aspects of a phenomenon to emphasize core causal mechanisms. In our context, idealizations help focus on the essential neural and biochemical pathways involved in pain without accounting for every physiological detail. This approach allows for a clearer understanding of the fundamental causal relationships defining pain in humans and octopi.

In humans, idealizing the neural pathways and biochemical mediators involved in pain processing enables us to highlight the core mechanisms, such as the role of nociceptors, afferent neurons, and specific brain regions like the thalamus and anterior cingulate cortex. Similarly, for octopi, idealizing their decentralized neural architecture and unique biochemical processes, such as the role of arm ganglia and neuropeptides, helps elucidate how pain is detected and managed without the need for a centralized brain.

Further applications and implications of integrating idealizations with the f_c mapping framework include enhancing our ability to create coherent and context-sensitive explanations of mental states across different beings. This approach not only prevents contradictions but also facilitates a deeper understanding of the causal mechanisms underlying pain and other mental states (Ryoo 2024a). By focusing on the core causal relations through idealizations, we can extend this framework to other areas of cognitive science, providing robust models for cross-species empathy and the ethical considerations of AI systems.

Appendix 1.1 Extended Analysis and Examples

Idealizations simplify complex systems by isolating key variables and processes, allowing us to focus on the most significant causal relationships. For instance, in the study of pain in humans, idealizing the role of specific neurotransmitters, such as serotonin and dopamine, can clarify their impact on pain modulation without delving into the intricate interactions with other neurotransmitters. This targeted approach aids in developing effective pain management therapies by emphasizing the most influential factors.

In the case of octopi, idealizing their response to injury by focusing on the primary neural and biochemical pathways involved in pain detection and response can reveal how these animals process pain differently from vertebrates. By highlighting the arm ganglia's role and the specific neuropeptides released during pain, we can better understand the evolutionary adaptations that enable octopi to survive in their unique environments.

Moreover, idealizations are crucial in the realm of artificial intelligence. When simulating pain in AI systems, we can idealize the computational processes by concentrating on the algorithms that detect damage and generate avoidance behaviors. This simplification allows researchers to refine these algorithms to improve AI responses to harmful stimuli, enhancing the robustness and reliability of AI systems in real-world applications.

Idealizations also play a vital role in educational contexts. By simplifying complex concepts, they make it easier to convey fundamental principles to students and non-experts. For example, using idealized models to teach the basic principles of pain perception and response can provide a clear and accessible introduction to the topic, laying the groundwork for more advanced and detailed studies.

The strategic use of idealizations within the f_c mapping framework facilitates a nuanced and effective approach to explaining mental states across different beings. By simplifying complex phenomena to focus on core causal mechanisms, idealizations enhance our understanding and ability to communicate these concepts. This method not only strengthens theoretical models but also supports practical applications in various fields, from developing AI systems to advancing medical treatments and promoting cross-species empathy. As we continue to refine and expand our use of idealizations, we can expect to gain deeper insights into the diverse manifestations of mental states and their underlying causes (Ryoo 2024b).

Appendix 2. Detailed Comparison of Functionalism Theories

In this section, we provide a detailed comparison of various forms of functionalism, including type functionalism, token functionalism, machine-state functionalism, psycho-functionalism, and analytic functionalism. We highlight their strengths and limitations and consider how the context-dependent mapping (f_c) framework could be applied to each.

Appendix 2.1 Type Functionalism

Type functionalism defines mental states by the general types of functional roles they play rather than their specific physical realizations. Mental states are characterized by their causal relationships with sensory inputs, behavioral outputs, and other mental states.

Strengths: Type functionalism accommodates multiple realizability, allowing the same mental state to be instantiated by different physical states across various organisms or systems. It provides a broad framework for understanding mental states across different contexts and entities.

Limitations: The high level of abstraction in type functionalism can sometimes overlook specific details of individual realizations. Additionally, determining the exact type of a mental state can be challenging due to the variability in realizations.

Mapping Implications: The f_c framework can enhance type functionalism by providing context-specific mappings that maintain the general functional roles while accommodating unique realizations. For example, mapping human pain to its neural, biochemical, and behavioral contexts ensures consistent explanations across different scenarios.

Appendix 2.2 Token Functionalism

Token functionalism emphasizes the specific instances (tokens) of mental states rather than the general types. Each occurrence of a mental state is considered in its unique physical realization.

Strengths: Token functionalism focuses on individual instances, providing detailed and specific explanations of mental states. It aligns closely with empirical observations and specific physical realizations.

Limitations: Token functionalism has difficulty in generalizing findings across different contexts or entities. Managing and explaining the multitude of specific instances can be complex and cumbersome.

Mapping Implications: The f_c framework can help token functionalism by organizing specific realizations into coherent subsets and similarity subspaces. This allows for detailed context-specific explanations while maintaining a structured approach to multiple realizations.

Appendix 2.3 Machine-State Functionalism

Machine-state functionalism, proposed by Hilary Putnam, likens mental states to the states of a Turing machine. Mental processes are viewed as computational processes that can be described by algorithms.

Strengths: Machine-state functionalism provides a clear and precise framework for understanding mental processes as computational states. It is particularly useful in the context of artificial intelligence and cognitive modeling.

Limitations: This form of functionalism can be overly reductive, potentially ignoring non-computational aspects of mental states. It may overlook the importance of biological and biochemical processes in natural organisms.

Mapping Implications: Applying the f_c framework to machine-state functionalism involves mapping computational states to specific contexts in which they operate. This helps integrate computational models with biological and biochemical considerations, providing a more comprehensive understanding of mental states.

Appendix 2.4 Psycho-Functionalism

Psycho-functionalism, advocated by Jerry Fodor, integrates psychological theories to define mental states by their roles in cognitive science. It emphasizes the importance of psychological laws and theories in explaining mental states.

Strengths: Psycho-functionalism directly connects mental states with psychological theories, making it highly relevant to cognitive science. It bridges the gap between philosophy of mind and empirical psychology.

Limitations: This form of functionalism relies heavily on the validity and acceptance of specific psychological theories. It may struggle with the variability of psychological states across different contexts and cultures.

Mapping Implications: The f_c framework can support psycho-functionalism by providing mappings that align psychological theories with specific contexts. This ensures that explanations remain consistent with psychological laws while adapting to different situational variables.

Appendix 2.5 Analytic Functionalism

Analytic functionalism, developed by David Lewis and others, uses common-sense psychological descriptions to explain mental states. It relies on the everyday understanding of mental terms and their causal roles.

Strengths: Analytic functionalism aligns with common-sense understandings of mental states, making it accessible and relatable. It can be applied across various contexts due to its reliance on general causal roles.

Limitations: The reliance on common-sense descriptions can lack the precision needed for rigorous scientific explanations. This form of functionalism can be vague and imprecise, leading to difficulties in defining and identifying specific mental states.

Mapping Implications: By applying the f_c framework, analytic functionalism can achieve greater precision and consistency. Context-specific mappings ensure that common-sense descriptions are grounded in concrete causal relationships, enhancing the robustness of explanations.

The f_c mapping framework offers significant advantages when applied to various forms of functionalism. By providing context-specific mappings and maintaining consistent causal roles, it enhances the explanatory power and coherence of each functionalist theory. This integration not only strengthens theoretical models but also supports practical applications in cognitive science, AI development, and the philosophy of mind.