How to Submit Proof Corrections Using Adobe Reader

Using Adobe Reader is the easiest way to submit your proposed amendments for your IGI Global proof. If you don't have Adobe Reader, you can download it for free at http://get.adobe.com/reader/. The comment functionality makes it simple for you, the contributor, to mark up the PDF. It also makes it simple for the IGI Global staff to understand exactly what you are requesting to ensure the most flawless end result possible.

Please note, however, that at this point in the process the only things you should be checking for are:

Spelling of Names and Affiliations, Accuracy of Chapter Titles and Subtitles, Figure/Table Accuracy, Minor Spelling Errors/Typos, Equation Display

As chapters should have been professionally copy edited and submitted in their final form, please remember that *no major changes to the text can be made at this stage*.

Here is a quick step-by-step guide on using the comment functionality in Adobe Reader to submit your changes.

1. Select the **Comment** bar at the top of page to View or Add Comments. This will open the **Annotations** toolbar.
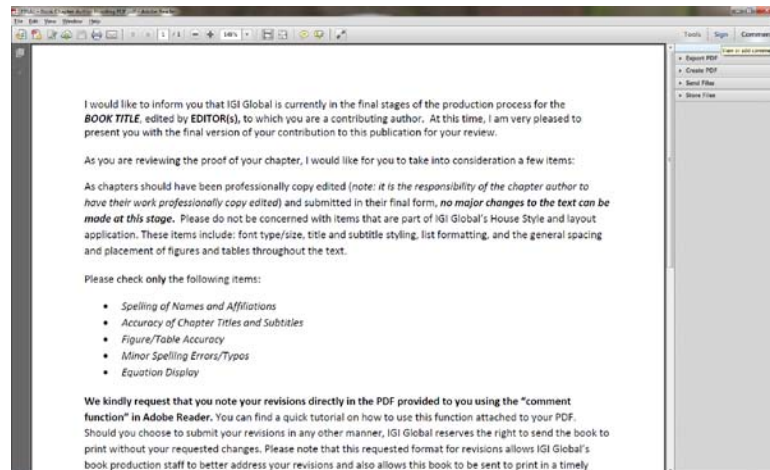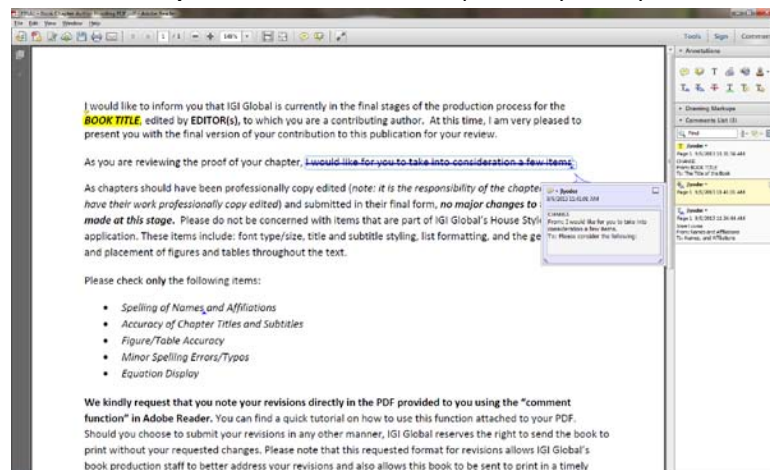

2. To note text that needs to be altered, like a subtitle or your affiliation, you may use the **Highlight Text** tool. Once the text is highlighted, right-click on the highlighted text and add your comment. Please be specific, and include what the text currently says and what you would like it to be changed to.
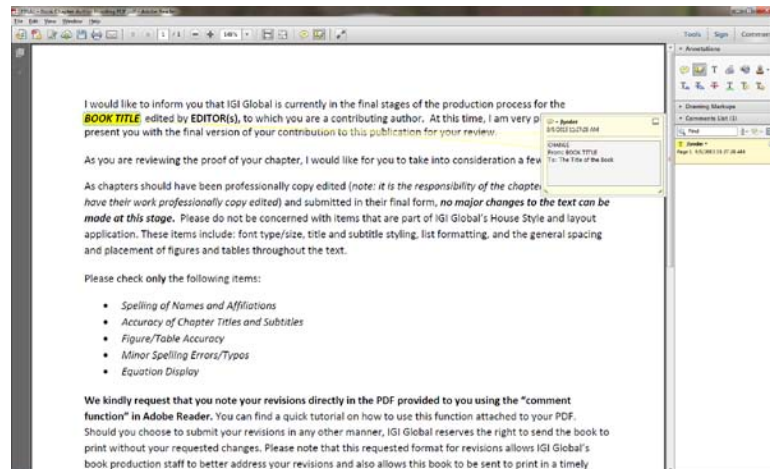
3. If you would like text inserted, like a missing coma or punctuation mark, please use the **Insert Text at Cursor** tool. Please make sure to include exactly what you want inserted in the comment box.
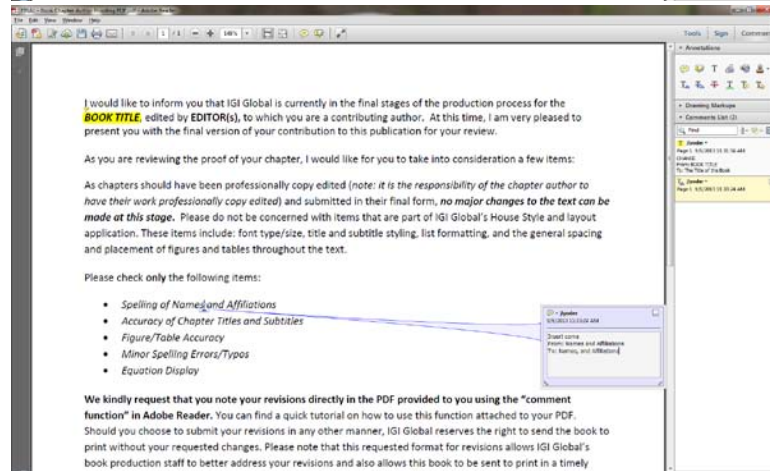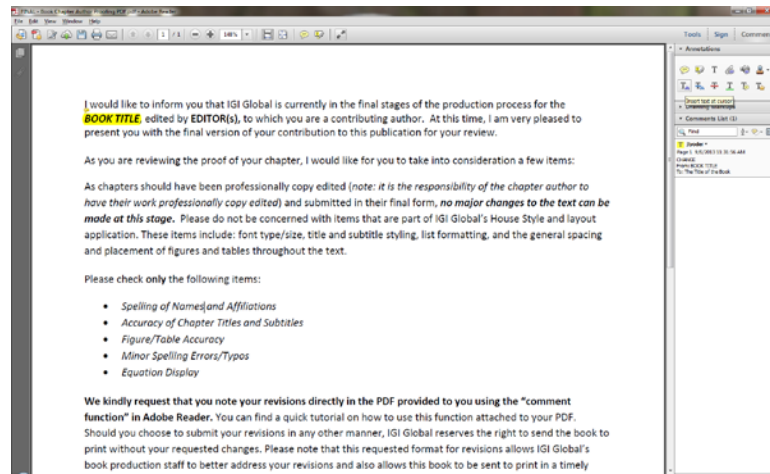
4. If you would like text removed, such as an erroneous duplicate word or punctuation mark, please use the **Add Note to Replace Text** tool and state specifically what you would like removed.

IGI PUBLISHING
WWW.IGI-GLOBAL.COM

Table of Contents

Robot Pain

Simon van Rysewyk, Graduate Institute of Medical Humanities, Taipei Medical University, Taipei City, Taiwan

ABSTRACT

Functionalism of robot pain claims that what is definitive of robot pain is functional role, defined as the causal relations pain has to noxious stimuli, behavior and other subjective states. Here, the author proposes that the only way to theorize role-functionalism of robot pain is in terms of type-identity theory. The author argues that what makes a state pain for a neuro-robot at a time is the functional role it has in the robot at the time, and this state is type identical to a specific circuit state. Support from an experimental study shows that if the neural network that controls a robot includes a specific 'emotion circuit', physical damage to the robot will cause the disposition to avoid movement, thereby enhancing fitness, compared to robots without the circuit. Thus, pain for a robot at a time is type identical to a specific circuit state.

Keywords: Emotion, Neuro-Robot, Pain, Role-Functionalism, Type-Identity Theory

INTRODUCTION

Is robot emotion possible? Currently, there are two main programs of scientific research investigating this question. One main line is a type of psychological behaviorism and assumes that emotion behavior can be explained in a robot without reference to subjective or mental events or states. According to this approach, the sources of robot emotion are exogenous (in the environment), not endogenous (in the head) (e.g., Adolphs, 2005; Canamero, 2005; Dautenhahn et al., 2009; Picard, 2000). Much work in this research program aims to build able robots to display emotional sensations behaviorally under specific circumstances without sensing them endogenously, or robots that attribute sensations to human beings based on visual observation of behavior rather than empathic understanding, conceived as an endogenous state type shared at the time of observation.

Behaviorist robot emotion faces several challenges. One challenge is that some subjective features of sensations have qualitative character, or 'qualia', broadly characterized as the properties of sensation states that type them in qualitative or sensed aspects (e.g., Chalmers, 1996; Jackson 1982; Tye, 2000). Type states of pain differ in sensed qualities from type states of pleasure, or in 'what it is like' to personally experience them. Thus, to have a pain is not only to produce appropriate pain behavior under certain circumstances, but it is to personally experience a 'like-this' sensory quality to the pain (e.g., as something burning or sharp). Two subjects in a psychophysical experiment presented with the same stimulus and showing the same behavioral responses may nonetheless have quite different qualitative experiences. Stimulus-response identity does not entail experiential identity, and such identity cannot be ensured other than by describing the experiential states and how these are experienced. In the same way, a purely

behaviorist robot may display pain behavior, yet completely lack pain qualia.

Some philosophers claim that what is definitive of individual sensation qualia ('quale') is functional role, described as the causal relations a sensory quale type has to stimuli, behavior and other subjective states (e.g., Armstrong, 1968; Fodor, 1975; Lewis, 1966; Maley & Piccinini, 2013; Putnam, 1975; Pylyshyn, 1984; Shoemaker, 1984). A role functionalist may define a pain as a sensory qualitative state that is reliably caused by noxious stimulation, to cause the desire to make it stop, to cause distraction regarding concurrent projects or plans and their completion, and to cause changes in preferences among alternative states of affairs. According to role-functionalism, only beings with subjective qualitative states that fulfill these causal roles can be in pain.

In normal adult humans, there appears to be a specific type of nervous-endocrine-immune biological activity that best fits these functional roles (van Rysewyk, 2013). Thus, normal humans can be in pain by noxious stimulation of the nervous-endocrine-immune ensemble. Since sensations that are physically very different may still feel the same, role-functionalism allows beings with different physical compositions to have sensory qualitative states as well. If there are biological states of feathered mites or non-biological states of robots that also fit these roles, then these beings can be in pain. Pain qualia can be realized by multiple types of physical states in multiple types of beings (e.g., Aizawa, 2008; Fodor, 1975; Putnam, 1975; Pylyshyn, 1984).

Role-functionalism is the current alternate program to psychological behaviorism in robot emotion research (e.g., Acerbi & Parisi, 2007; Breazeal & Brooks, 2005; Parisi & Petrosino, 2010; Pérez et al., 2012; Ziemke, 2008). In this article, I propose that the results of Parisi & Petrosino (2010) and philosophical theory both support my claim that what is definitive of robot pain and emotion generally is not simply functional role, as interpreted by role functionalists, but a tandem product of functional role and type-type identity; that is, what makes a state

pain for a robot at a time is the functional role it has in the robot at the time, and this state is identical to a state of a specific type of circuit added to a robot's neural network. In the next section, I describe in detail the experiment conducted by Parisi & Petrosino (2010), including its theory, methods, and results related to pain.

Role-Functional Robot Pain

Parisi & Petrosino (2010) test the hypothesis that functional role is definitive of robot emotion, not behavior. The authors introduce two further sub-hypotheses: (1) robot emotion requires robots to have different motivations at a time that compete with one another for control of the robot's behavior; (2) robot emotion requires the neural network that controls the robot's behavior to have a circuit which engages the robot to take faster and more correct motivational decisions. The study involved five different simulated Khepera robots which were required to satisfy two different motivations: (1) eat and drink, (2) eat and avoid a predator, (3) eat and find a mate, (4) eat and care for their offspring, or (5) eat and rest in order to recover from physical damage and pain. The authors predicted that adding the emotional circuit to the neural network will lead to better motivational decisions and higher fitness in each simulation, compared to robots without the circuit. I will focus in this paper on the pain robots ('5').

Sub-Hypothesis 1: Emotion Changes the Disposition of Motivation

Imagine a monkey perceiving a thorny berry bush. The monkey approaches the bush, picks berries and eats them. What causes this event? According to psychological behaviorism, exogenous sensory stimuli and an organism's interactions with, and reinforcement from, the environment cause this event type (e.g., Rescorla, 1990; Skinner, 1953). Parisi & Petrosino (2010) suggest this explanation is half right, since a behavioral change also implies certain things about an animal's motivations, not only sensory stimuli. After all, the monkey likely

approaches the bush because of hunger. For example, the monkey probably wants the hunger sensation to stop. If the monkey is satiated, it will behave indifferently toward the bush, and likely disregard it. So, the question is: can the mere presence of an environmental sensory stimulus imply anything about the monkey's wants and motivations? Parisi & Petrosino (2010) answer in the negative because it is not clear how there can be a *sensation* of *wanting* to avoid a hunger stimulus, or of *wanting* to eat. Instead, it appears that animal behavior is the joint effect of environmental sensory impingement and a change in the disposition of an animal's motivations. It follows from this view that what is definitive of hunger is a specific type of sensory qualitative state which causes a compelling motivation that such sensing cease.

The same argument effectively applies to understanding the role of sensory qualia in emotion. Consider two separate pain events in an animal. Suppose that these events are qualitatively identical, that share all the *same* sensory pain qualia, yet which are *not* equally painful. How? If the same pain state is clustered with different wants and motivations, this second pain state may not be equally painful. If Parisi & Petrosino (2010) are right, an emotion sensation implies motivational change. But, as the present imagined case shows, the mere presence of a quale cannot imply anything about the animal's motivations. It follows that the mere presence of a quale cannot guarantee any particular effect on motivational state. That is, no emotion quale is itself emotional because it alone cannot guarantee any motivational difference. In the event of pain, a pain sensation must arouse the disposition to avoid before the effect is 'pain' (Clark, 2005). Like the role functional definition of hunger suggested above, what identifies pain is a particular type of sensory qualitative state which causes an immediate and compelling motivation to avoid such sensing.

Accordingly, the way to effect robot emotion is to wire the perception of sensory qualities directly into the robot's motivational functions. Sensing in that way will cause an immediate motivational change. Parisi and Petrosino (2010) assume a simple mechanism for controlling which specific motivation will cause the robot's behavior at any given time: all the robot's motivations have a quantitative level of intensity and the motivation which wins the competition is the motivation which currently has the highest level of intensity (Parisi, 2004; Ruini et al., 2010). The current level of intensity of a motivation may depend on a number of factors such as the particular environment in which the robot has evolved, including environmental stimuli, the role of the motivation in the robot's overall adaptive pattern, and the state of the robot's body.

The pain robots are in an environment of 1000 x1000 pixels with four food patches of 40 pixel diameters which they eat in order to survive. They nonsexually reproduce at regular intervals. At random intervals, the body of these robots may receive some physical damage which can vary in intensity and which persists for an irregular time. In the event of pain, the robots must rest or reduce their speed in order to heal, even when very hungry. If they do not rest or reduce movement speed due to pain, their fitness is proportionally reduced. If they decide to search for nutrients, they have to move very quickly, despite pain. The robot's life is five epochs, each epoch a maximum of 2000 time-steps.

Artificial Neural Network in the Robots

The robots in Parisi and Petrosino (2010) are controlled by an artificial neural network. The architecture of the neural networks consists of input units representing information from the exogenous environment (food, water, predator, mating partner, offspring-care zone), from the endogenous environment (hunger, thirst, pain), internal units, and motor output units which encode the speed of the robots wheels (Figure 1). In contrast to classical symbolic theories of mental processing, information such as 'pain', 'motivation', and 'behavior', is stored non-symbolically in the weights, or connection strengths, between the units in neural network,

Figure 1. Artificial neural network architecture for control-group robots no emotional circuit (Parisi & Petrosino, 2010)

OUTPUT LAYER

INTERNAL LAYER

INPUT LAYER
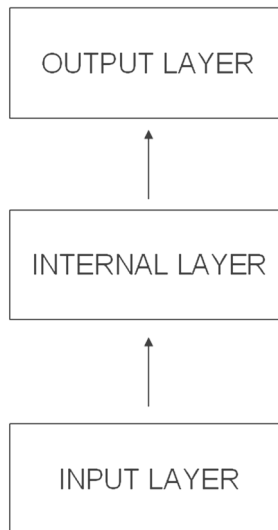
according to the activation function. The neural network in the study is a standard network with activation level between 0 and 1 for all the network's units and a sigmoid activation function for the internal units and the motor units.

The neural network in the pain robots is informed of physical damage and of its severity by a continuously activated pain sensor. Reduction in fitness due to physical damage is in proportion to damage severity and to how the robot responds to pain (e.g., by reducing its movement speed). When a robot senses pain, the sensing will cause a compelling and specific motivational change: either the robot ceases all movement, or the robot reduces movement speed to search for nutrients. In the study, a pain sensation was characterized to cause the disposition to minimize or avoid all movement before the result is 'pain'.

Sub-Hypothesis 2: Emotion Requires an Emotion Circuit

To test the study hypothesis that functional role is definitive of robot emotion, Parisi & Petrosino (2010) compared two different populations of

robots for each of the five types of robots. The control population of robots in each type had the neural network described above. To the neural network in the experimental population was added an 'emotional circuit', which consists of one or two 'emotional units' to which some of the input units send their activation. These units send their activation either to the internal units (Figure 2a) or to the motor units (Figure 2b).

The emotional unit in the experimental pain robots receives activation from the pain sensor which informs the robots to stop moving when there is pain. The pain sensor has an activation of 1 for maximum physical damage/maximum pain and an activation of 0 for zero physical damage/no pain. The activation threshold of the emotional unit is represented in the robot's inherited genotype together with the connection weights of the robot's neural network. The authors expected that the emotional circuit will arouse a specific motivational effect such that an experimental pain robot with the circuit will avoid moving in the event of physical damage concurrent with hunger, thereby enhancing personal fitness. Thus, concurrent and equally

severe pain and hunger should cause pain robots with the emotional circuit to avoid all movement, even when nutrients are available, as per the study hypothesis.

Pain Robot Results

Comparison of the two pain robot groups shows that pain experimental robots with an emotional circuit had higher fitness (Figure 3) than pain control robots without an emotional circuit (Figure 4).

The experimental pain robots moved only when pain was very low, and proportional to food distance. That is, corresponding pain and hunger matched in severity reduced average movement velocity in all pain robots relative to nutrient distance. When corresponding pain and hunger were each severe in pain robots with the emotional circuit linked with the motor units, they avoided all movement (Figure 5). *Per* hypothesis, painfulness must arouse an aversion before the result is 'pain' (Clark, 2005).

Parisi and Petrosino (2010) interpret these findings to imply that the pain robots with an emotional circuit have pain states because these states are the activation states of their emotional circuit and these emotional states have a functional role in their behavior (e.g., they avoided moving entirely when disposed

to seek nutrients). Linking the perception of sensory qualities into the robot's motivational functions causes a change in the disposition of a robot's motivation. I claim the only way to understand this interpretation is in terms of type-identity theory, a philosophy of mind. In the next section, I introduce mind-brain identity theory, and type-identity theory.

Mind-Brain Identity Theory

Mind-brain identity theory asserts that what is definitive of sensation is that sensation things (e.g., states, processes, properties) are *identical* with brain things (e.g., states, processes, properties), not merely correlated with them, caused by them, realized by them, or constituted by them (e.g., states, processes, properties) (e.g., Churchland, 1989; Feigl, 1958; Hill, 1991; Place, 1956; Smart, 1959; van Rysewyk, 2013).

There are two main versions of the mind-brain identity theory: token identity theory and type-identity theory. Token identity proposes that individual sensation states, or *tokens* of sensation states, are identical with individual brain states, or *tokens* of brain states. For example, the pain I currently feel in my left thigh is identical with a particular state of my brain. The other main version of the identity theory asserts that an individual sensation state is a

Figure 2. Artificial neural network architecture for experimental-group robots with an emotional circuit connected to (b) the internal layer or (2) to the output layer (Parisi & Petrosino, 2010)
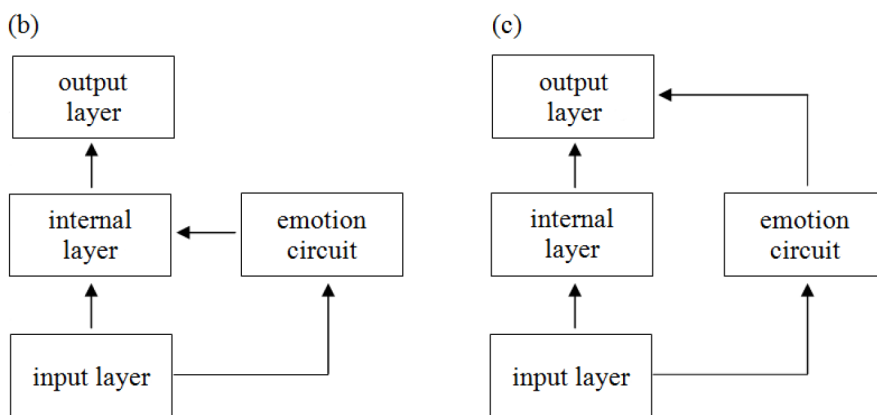
Figure 3. Best and average fitness in pain experimental robots (Parisi & Petrosino, 2010)
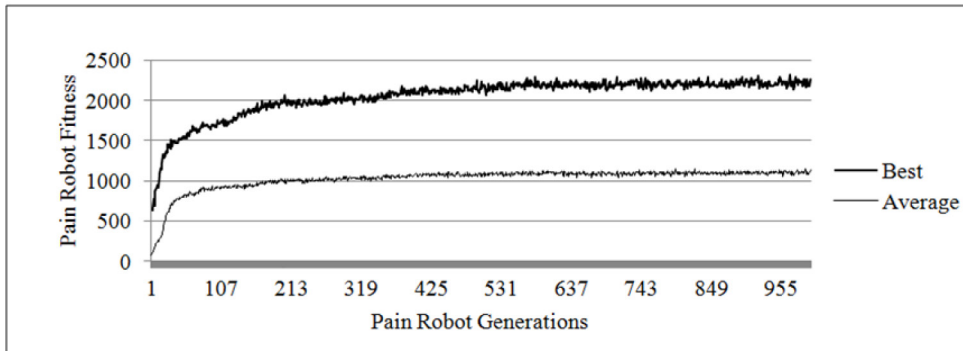
Figure 4. Best and average fitness in pain control robots (Parisi & Petrosino, 2010)
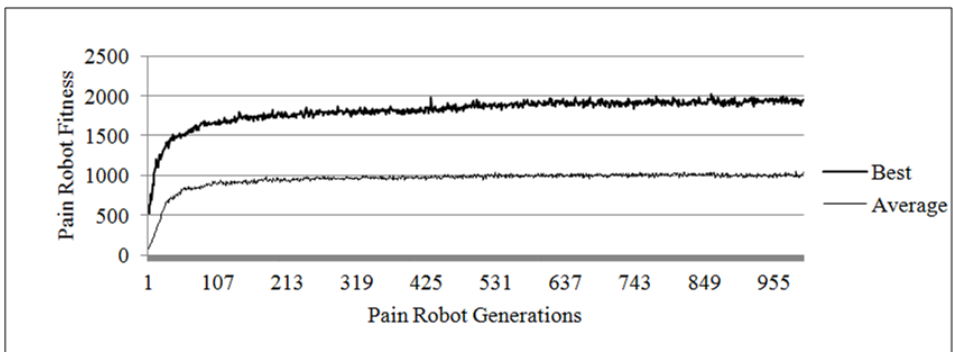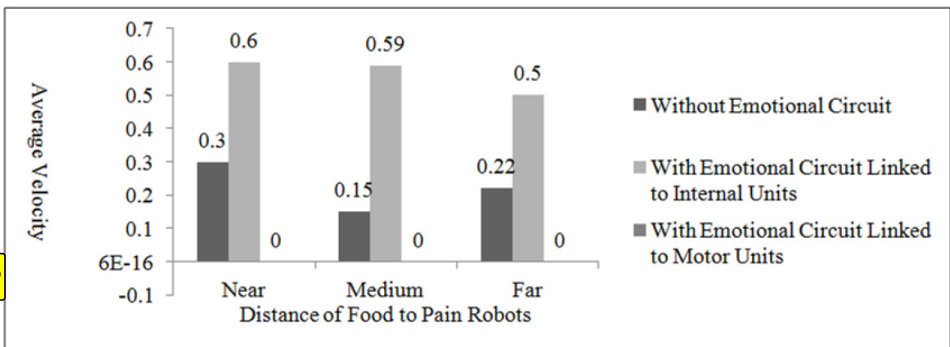
Figure 5. Average movement velocity of very hungry pain experimental robots when food was near, at medium distance or far away (Parisi & Petrosino, 2010)

sensation *type* identical with a brain state *type*. So, in addition to implying that my current pain is identical with a specific brain state, it implies that the state being a pain is identical with a type of brain state. This means that assertions of type identity are logically stronger than assertions of token identity, and have accordingly greater explanatory potential to offer an understanding of sensation state-brain state relations.

In what follows, I focus on type-identity theory, and accordingly view 'pain' and 'state' as naming universal types, which may occur in the same creature at different times, or which many different creatures may share, and not as naming particular tokens, which may only occur exactly once in the world (if any) (Jackson, 2012).

To illustrate type identity with an actual case, assume two subjective type states, S1 and S2, robustly connected, respectively, to two neurobiological type states, P1 and P2. Thus, some cutaneous wounds initially generate a highly localized 'first' pain (S1) that is followed by a poorly localized 'second' pain (e.g., stubbing a toe) (S2). This experience is called 'double pain sensation' (Campbell & LaMotte, 1983). The temporal ordering of S1 and S2 is robustly connected with nociception: firing of Type II Aδ nociceptors leads to the 'first' acute pain response to noxious heat (P1). Compression block of myelinated peripheral nerve fibers eliminates first, but not second, pain (Torebjörk & Hallin, 1973). The first of the double pain sensations reaches the central nervous system (CNS) on the neospinothalamic tract and activates the ventroposterolateral and ventroposteromedial nuclei of the thalamus and then primary somatosensory cortex (Willis & Westlund, 1997) (P1). Alternately, a low dose of local anesthesia applied to peripheral nerves blocks the C-fibers before the Aδ-fibers (Johansson et al., 1990). Under this condition, the slow conducting pain information is blocked (paleospinothalamic tract), and only the fast conducting pain information by Aδ-fibers is carried to the CNS. The second of the double pain sensations (S2) reaches the CNS on the paleospinothalamic tract to activate brainstem nuclei and then the parafasciculus and centro-median mechanism in the intralaminar thalamic nuclei (Willis & Westlund, 1997) (P2).

Type-identity theory denies the claim made by property dualism and epiphenomenalism that though pains are brain states they still have fundamentally nonphysical, psychical properties (e.g., Chalmers, 1996; Jackson, 1982). Briefly, there are five reasons to prefer type identity to dualist theories (van Rysewyk, 2013). The first reason is that type identity possesses greater *explanatory power* (Place, 1956). Type-identity theory observes the robust empirical correlations between subjective states and specific neurobiological states and between subjective states and specific behavioral states (e.g., facial expressions) reported in the neurosciences and psychology. Thus, when a functional neuroimaging study strongly correlates a cognitive operation with a brain region (e.g., perceiving human faces with cortical activity in the fusiform face area), or when an event-related potential study identifies a cognitive operation with a change in electrical potentials (e.g., age-differences in human face perception correlated with a N170 event-related potential), type-identity theory claims that the neural operation *is* the cognitive operation: increased activity in the FFA *is* the visual perception of human faces; increased activity in the N170 *is* the age-difference in human face perception.

Type-identity theory asserts that it offers the best description and the best explanation for these robust correlations. That is, the best way of describing and explaining the robust correlation between pain qualia and specific neurobiological states is to assert that pain is type identical with those states. According to a kind of abductive inference called *Inference to the Best Explanation*, it is appropriate to prefer descriptive theories that offer the best explanations of phenomena in their domains, all other things being equal. Thus, if a theory provides the best explanation of all the data that are relevant to pain, then we are allowed to believe that descriptive theory, relative to the alternatives (e.g., dualism). Since type identities are thought by type identity theorists to satisfy

this stipulation much better than dualism, type identity philosophers infer that type-identity theory best describes the mind.

The second reason to prefer identity theory is that type identity is the only theory that fully respects philosophical intuitions about the causal powers of qualia. For example, we believe that pains are causally responsible for such behaviors as facial grimaces, limb- guarding, crying out, and also for much of our talk and thought about pain. We attribute important causal powers to pain, but neuroscience indicates that these causal powers also reside in the biological operation that is type identified with pain. Accordingly, if pain is distinct from that biological operation, we will be required to infer either that pain is epiphenomenal, having no causal powers in its own right, but merely appearing to have them because of its relation with the neurobiological state (Jackson, 1982). In itself, epiphenomenalism is not an attractive view. It obliges us to believe that subjective states, even though they are caused by neurobiological operations, have *no effects* on the world. This seems a very strange kind of causal power. For example, if pains don't *cause* pain behavior how can it be that your telling me that you are in pain gives me any reason for supposing you are? Moreover, if pain is absolutely epiphenomenal, then a search for fundamental type identities of pain will fail. In fact, if pain is completely epiphenomenal, then it cannot have evolved by natural selection.

Third, type identity sees only *one category* of states where other mind philosophies see two. The view of reality that type-identity theory offers is simpler, and more coherent, than the view that is offered by property dualism, which proposes that qualia are ontologically unique states, having no descriptive and explanatory role in physics, biology, or any other natural science. Whereas type-identity theory sees only a single state – a quale that is a neurobiological state – property dualism sees two. All other things being equal, I think it is reasonable to say that simpler theories are preferable to complex ones.

Fourth, type-identity theory is supported by the causal closure thesis, according to which every physical effect has an immediate sufficient physical cause (e.g., Papineau, 2009). This thesis comprises three requirements: every physical effect has a cause which is *physical*, *immediate*, and *sufficient*. The first requirement says that for any physical effect, there will always be a prior physical cause. The second requirement that the physical cause be immediate is needed to exclude the possibility of physical causes which produce their physical effects only via nonphysical intermediaries. Finally, the third requirement that the physical cause be sufficient is needed to establish that it causes the physical effect by itself and not solely in virtue of its conjunction with some ontologically unique nonphysical cause such as a nonphysical quale. While the causal closure of the physical lies below the surface in type-identity theory, causal closure plays an essential descriptive and explanatory role in it: if causal closure were not true, then some physical effects would not be determined by prior physical causes at all (e.g., brain states), but by ontologically unique subjective causes such as nonphysical qualia.

Finally, type-identity theory can be translated into psychometric terms and treated like a hypothesis that can be empirically assessed and compared with token identity theory and other philosophies of mind. For example, Kievit et al. (2011) formalize type-identity theory using a measurement model called an effect indicator model in which a hypothetical but not directly observable concept is estimated by observed, measurable subjective and neural variables. In this model, the measured subjective and neural variables correlate with each other perfectly because they have a common cause: the hypothetical concept. As I hope to show in this paper, robots are yet another way of formulating and empirically testing type-identity theory (Parisi, 2010; Parisi & Petrosino, 2010).

Role-Functionalism as a Route to Mind-Brain Identity Theory

With type-identity theory introduced, I will now present an argument to theorize role-functionalism of robot pain in terms of type-identity theory. Recall that role-functionalism defines

sensations in terms of functional role, characterized as the causal relations a sensation has to stimuli, behavior and other subjective states. Thus, what makes a state pain for a creature at a given time is that it has the pain-functional role in the creature at the time; in biological creatures, likely a brain state involving interdependent nervous, endocrine and immune activity (van Rysewyk, 2013). From that sentence of two clauses we obtain the following two premises, respectively:

1. What makes a state pain for a creature at a given time is that it has the pain-functional role in the creature at the time;
2. The state which has the pain-functional role for a creature at a given time is a brain state involving interdependent nervous, endocrine and immune activity;

From these premises in conjunction with the prior stipulation that 'pain' and 'state' name types, not individual tokens (Jackson, 2012), I infer by transitivity the following conclusion:

3. Therefore, a state pain for a creature at a given time is type identical to a brain state involving interdependent nervous, endocrine and immune activity.

This conclusion presents a type-identity theory and not a token identity theory because in 'A state pain for a creature at a given time = a brain state involving interdependent nervous, endocrine and immune activity', which is a type-type identity; 'a brain state involving interdependent nervous, endocrine and immune activity' describes a type of event, which holds irrespective of whether any creature is currently in pain. For a creature at a given time to be in pain is to have a particular instance of the *type* of state in question at the time.

In the same way, the behavior of the experimental pain robots observed in Parisi & Petrosino (2010) presents a type-type identity. The philosophical rationale of their study can be conveyed in the form of my argument just stated; namely:

1. What makes a state pain for a robot at a given time is that it has the pain-functional role in the robot at the time;
2. The state which has the pain-functional role for a robot at a given time is a neural network state involving activity of an emotional circuit;
3. Therefore, a state pain for a robot at a given time is type identical to a neural network state involving activity of an emotional circuit.

As in conclusion 3, the inference 3* introduces a type-identity theory and not a token identity theory because in 'A state pain for a robot at a given time = a neural network state involving activity of an emotional circuit', which is a type-type identity, 'a neural network state involving activity of an emotional circuit' refers to an event type. Even if the emotional circuit linked to internal units was redesigned at a time during the experiment to instead link with motor units, pain robot states before and after that time would be, if true, type-type identities, since a specific robot-design is a type, not a token. To claim that pain for a creature at a given time is what has the pain functional-role for the creature at the time is not to claim that the creature is in pain at the time. Again, it is only to claim that for the creature to be in pain at a time is to have an instance of the *type* of state in question at the time. When roboticists such as Parisi & Petrosino (2010) create the design formulas for pain in a robot, they are creating for a type. The same logic applies to robot sensation and emotion generally.

My interpretation of Parisi & Petrosino (2010) is strongly supported in the conclusion of the entire study in their own words: 'The robots endowed with an emotional circuit can be said to have emotions or emotional states. *Their emotional states are the activation states of their emotional circuit* and these emotional states have a functional (beneficial) role in their behavior' (Parisi & Petrosino, 2010, p. 9, my italics). Thus, the argument stated in Parisi & Petrosino (2010) can only be understood as implying a type-identity theory.

CONCLUSION

This paper addressed robot pain to exemplify robot emotion. I argued that role-functionalists concerning robot emotion and sensation can and should make an additional type-type identity claim that subjective states are brain states. Making that extra claim means that what makes a state pain for a creature at a given time is that it has the pain-functional role in the creature at the time; in biological creatures or non-biological systems such as robots, what is type identical to this functional-role is a brain state.

To create pain robots it is necessary to work with machines that have more than one dispositional motivation; competition between the motivations must result in an overriding aversion, or some disposition to avoid, if the result is to be painfulness. The emotional circuit succeeded in arousing a compelling aversion, despite concurrent hunger. Thus, the pain robots in Parisi & Petrosino (2010) have pain because it is possible to describe the functional role that pain states have in their behavior and to type-identify the specific part of the neural network causing their behavior that makes pain states possible.

Neuro-robots can be used to empirically test philosophies of mind and emotion because they permit researchers to examine the representations in the robot's brain (i.e., artificial neural network) to establish if the behavior of the robot and its organs and systems are due to sensory input rather than sensory input *per se*. The results of these examinations can be integrated with ecological and evolutionary data on organisms, especially data of brain-body relationships, the results of behavioral and neuropsychological experiments, and philosophical theory.

This novel suggestion could enable a powerful functionalist-type-identity-based research program, productively extending the traditional brain-focus of type-identity theory to address the broader phenomena of embodiment, and concepts in Artificial Intelligence such as morphological computation. An important test of this program will be its ability to reproduce not only the results of experiments and psycho-metric models but also other empirical facts such as psychiatric and psychological disorders, intra- and inter-individual differences in emotion and pain, and social phenomena, such as pain empathy.

ACKNOWLEDGMENT

REFERENCES

Acerbi, A., & Parisi, D. (2007). The evolution of pain. In W. B. T. Christaller, P. D. J. T. Kim, & J. Ziegler (Eds.), *Advances in artificial life* (pp. 816–824). Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-74913-4_82

Adolphs, R. (2005). Could a robot have emotions? Theoretical perspectives from social cognitive neuroscience. In J. M. Fellous, & M. A. Arbib (Eds.), *Who needs emotions? The brain meets the robot* (pp. 9–29). Oxford, UK: Oxford University Press.

Aizwa, K. (2008). Neuroscience and multiple realization: A reply to bechtel and mundale. *Synthese*, *167*, 495–510.

Armstrong, D. (1968). *A materialistic theory of the mind*. London, UK: Routledge.

Breazeal, C., & Brooks, R. (2005). Robot emotion. A functional perspective. In J. M. Fellous, & M. A. Arbib (Eds.), *Who needs emotions? The brain meets the robot* (pp. 271–310). Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780195166194.003.0010

Campbell, J. N., & LaMotte, R. H. (1983). Latency to detection of first pain. *Brain Research*, *266*(2), 203–208. doi:10.1016/0006-8993(83)90650-9 PMID:6871657

Canamero, L. (2005). Emotion understanding from the perspective of autonomous robots research. *Neural Networks*, *4*, 445–455. doi:10.1016/j.neunet.2005.03.003 PMID:15963689

Chalmers, D. (1996). *The conscious mind*. Oxford, UK: Oxford University Press.

Churchland, P. S. (1989). *Neurophilosophy: Toward a unified science of the mind-brain*. Cambridge, MA: MIT Press.

Clark, A. (2005). Painfulness is not a quale. In M. Aydede (Ed.), *Pain: New essays on its nature and the methodology of its study* (pp. 177–198). Cambridge, MA: MIT Press.

Dautenhahn, K., Nehaniv, C. L., Walters, M. L., Robins, B., Kose-Bagci, H., Assif Mirza, N., & Blow, M. (2009). KASPAR: A minimally expressive humanoid robot for human–robot interaction research. *Applied Bionics and Biomechanics*, *6*, 369–397. doi:10.1080/11762320903123567

Feigl, H. (1958). The 'mental' and the 'physical'. *Minnesota Studies in the Philosophy of Science*, *2*, 370–497.

Fodor, J. (1975). *The language of thought*. New York, NY: Thomas Crowell.

Hill, C. S. (1991). *Sensations: A defense of type materialism*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9781139173827

Jackson, F. (1982). Epiphenomenal qualia. *The Philosophical Quarterly*, *32*, 127–136. doi:10.2307/2960077

Jackson, F. (2012). In defence of the identity theory mark I. In S. Gozzano, & C. S. Hill (Eds.), *New perspectives on type identity: The mental and the physical* (pp. 150–166). Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511687068.008

Johansson, A., Hao, J., & Sjölund, B. (1990). Local corticosteroid application blocks transmission in normal nociceptive C-fibres. *Acta Anaesthesiologica Scandinavica*, *34*(5), 335–338. doi:10.1111/j.1399-6576.1990.tb03097.x PMID:2167604

Kievit, R. A., Romeijn, J. W., Waldorp, L. J., Wicherts, J. M., Scholte, H. S., & Borsboom, D. (2011). Mind the gap: A psychometric approach to the reduction problem. *Psychological Inquiry*, *22*(2), 67–87. doi:10.1080/1047840X.2011.550181

Lewis, D. (1966). An argument for the identity theory. *The Journal of Philosophy*, *63*, 17–25. doi:10.2307/2024524

Maley, C., & Piccinini, G. (2013). Get the latest upgrade: Functionalism 6.3. 1. *Philosophia Scientiæ*, *17*(2), 135–149. doi:10.4000/philosophiascientiae.861

Papineau, D. (2009). The causal closure of the physical and naturalism. In B. McLaughlin, A. Beckermann, & S. Walter (Eds.), *The Oxford handbook of philosophy of mind* (pp. 53–65). Oxford, UK: Oxford University Press. doi:10.1093/oxfordhb/9780199262618.003.0003

Parisi, D. (2004). External robotics. *Connection Science*, *16*, 325–338. doi:10.1080/09540090412331314768

Parisi, D. (2011). The other half of the embodied mind. *Frontiers in Psychology*, *2*, 1–8. doi:10.3389/fpsyg.2011.00069 PMID:21713130

Parisi, D., & Petrosino, G. (2010). Robots that have emotions. *Adaptive Behavior*, *18*(6), 453–469. doi:10.1177/1059712310388528

Pérez, C. H., Escribano, G. S., & Sanz, R. (2012). The morphofunctional approach to emotion modelling in robotics. *Adaptive Behavior*, *20*(5), 388–404. doi:10.1177/1059712312451604

Picard, R. W. (2000). *Affective computing*. Cambridge, MA: MIT Press.

Place, U. T. (1956). Is consciousness a brain process? *The British Journal of Psychology*, *47*, 44–50. doi:10.1111/j.2044-8295.1956.tb00560.x PMID:13304279

Putnam, H. (1975). *Mind, language, and reality*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511625251

Pylyshyn, Z. (1984). *Computation and cognition*. Cambridge, MA: MIT Press.

Rescorla, R. A. (1990). The role of information about the response-outcome relationship in instrumental discrimination learning. *Journal of Experimental Psychology. Animal Behavior Processes*, *16*, 262–270. doi:10.1037/0097-7403.16.3.262 PMID:2398337

Ruini, F., Petrosino, G., Saglimbeni, F., & Parisi, D. (2010). The strategic level and the tactical level of behaviour. In J. Gray & S. Nefti-Meziani (Eds.), Advances in cognitive systems (pp. 271-299). Herts: IET Publisher.

Shoemaker, S. (1984). *Identity, cause, and mind*. Cambridge, UK: Cambridge University Press.

Skinner, B. F. (1953). *Science and human behavior*. New York, NY: Macmillan.

Smart, J. J. C. (1959). Sensations and brain processes. *The Philosophical Review*, *68*, 141–156. doi:10.2307/2182164

Torebjörk, H. E., & Hallin, R. G. (1973). Perceptual changes accompanying controlled preferential blocking of A and C fibre responses in intact human skin nerves. *Experimental Brain Research*, *16*(3), 321–332. doi:10.1007/BF00233334 PMID:4686614

Tye, M. (2000). *Consciousness, color, and content*. Cambridge, MA: MIT Press.

van Rysewyk, S. (2013). *Pain is mechanism*. Doctoral dissertation, University of Tasmania.

Willis, W. D., & Westlund, K. N. (1997). Neuro-anatomy of the pain system and of the pathways that modulate pain. *Journal of Clinical Neurophysiology*, *14*, 2–31. doi:10.1097/00004691-199701000-00002 PMID:9013357

Ziemke, T. (2008). On the role of emotion in biological and robotic autonomy. *Bio Systems*, *91*, 410–408. doi:10.1016/j.biosystems.2007.05.015 PMID:17714857