

Emotion Researcher

ISRE's Sourcebook for Research on Emotion and Affect

<http://emotionresearcher.com>

July 2017

Editors

Carolyn Price

Eric A. Walle

Open University
carolyn.price@open.ac.uk

University of California, Merced
ewalle@ucmerced.edu

Table of Contents

Editors' Column <i>Carolyn Price & Eric A. Walle</i>	3 – 4
ISRE Matters <i>Arvid Kappas</i>	5 – 8
ISRE Biennial Meeting Update <i>Hillary Anger Elfenbein</i>	9
ISRE Biennial Meeting Update <i>W. Gerrod Parrott</i>	10
In Memoriam: Jaak Panksepp <i>Douglas F. Watt</i>	11 - 19
ISRE Interview: The Social Construction of Emotion: Myths and Realities <i>James Averill</i>	20 - 43
ISRE Spotlight: Four Unwarranted Assumptions about the Role of Emotion in Moral Judgement <i>Nina Strohminger</i>	44 - 52
Feature Articles: Empathy	
Recent Work in the Philosophy of Empathy <i>Derek Matravers</i>	53 - 58
Empathy and its Development: What is Missing? <i>Alexandra Main</i>	59 - 65
The Therapeutic Effects of Empathy in Healthcare <i>Jodi Halpern</i>	66 - 71

Editors' Column

Emotion Researcher: Past, Present, Future

Carolyn Price & Eric A. Walle

We are excited to take over the editorship of Emotion Researcher. As we take on the associated responsibilities, we felt it important to reflect on the development of this publication, reaffirm those aspects that we feel are most central to its mission, and highlight some directions that we aim to move toward in the coming years.

Past

Emotion Researcher was originally developed to serve as a newsletter for the International Society for Research on Emotion. The goal was to provide the ISRE membership with a shared source of information for items relating to the society and the broader field.

Early editors included Randolph Cornelius, Agneta Fischer, Kim Bard, and Nathan Consedine. More recently the publication was edited by Christine Harris, who recalls spending hours (sometimes 40+ in a given week) formatting each issue, the paper copy of which necessitated that the newsletter consist of a multiple of 4 pages due to the folding of the sheet!

Andrea Scarantino took over as Editor in 2013. Under his guidance over the past 3 years, Emotion Researcher has progressed from a regular newsletter to a permanent, expanding, online resource where anyone who is interested in emotion can access expert, cross-disciplinary discussions concerning fundamental aspects of emotion. It is noteworthy that the efforts put into developing Emotion Researcher by Editors was, and continues to be, completely *pro bono* and with little or no editorial assistance.

Present

As we take over as Co-Editors of Emotion Researcher, we will work to ensure that the fundamental tenement of this publication continues in every sense of the word, namely *accessibility*.

First, we are a completely online and openly accessible publication. One does not need a membership or subscription to access the content; simply an interest in emotion. This allows people from around the world, regardless of class or status, to have access to cutting edge research in the field of emotion.

Second, there is an accessibility of ideas to individuals from disparate backgrounds and fields of study. This embodies the spirit of ISRE as an inclusive enterprise that brings together scientists, researchers working in the natural sciences, social sciences, and humanities. The articles are written to be interpretable by experts and non-experts alike. This makes the articles useful for anyone interested in the topic of emotion, a topic that is inherently complex and can be studied from a wide variety of perspectives. Additionally, the articles are interpretable for individuals just getting into the field of emotion research or anyone seeking insights from a discipline outside of their own.

Finally, Emotion Researcher has a personal quality that makes the people in the field accessible. Academia can feel stifling and impersonal, particularly for younger researchers. The interviews and spotlight features of Emotion Researcher are one means for alleviating this issue. Whether it was sharing the recipe for Nico Frijda's homemade mango soup or allowing us to see a young Jennifer Lerner's "enthusiastic eating," this publication makes the people "really real" for the reader. Whereas most academic publications seek to filter out and control the personal, "unscientific" aspects of the author, we will continue to share the person as much as the ideas.

Future

In addition to continuing the aims mentioned above, we also have some new initiatives that we hope to bring to the Emotion Researcher.

First, we will strive to increase our visibility. This task will begin in earnest at the Biennial ISRE Meeting in St. Louis. As the outgoing Editor, Andrea will take part in the Meet the Editors Workshop on July 26th. This meeting will bring together the editors from the top outlets in the field of emotion and we are thrilled that Emotion Researcher will be included. We will also work with the incoming ISRE Board to

ensure that interested members receive each issue.

Second, we will work to raise the reputation of Emotion Researcher. There are plenty – some might argue too many – journals already being published. Rather than making Emotion Researcher another addition to the pile, we hope to carve out a unique niche. Our existing online presence will help facilitate this goal, allowing researchers to access comprehensive and cross-disciplinary perspectives on fundamental topics in emotion, as well as providing flexibility to try out new ideas and features. Relatedly, you will notice that the distributed “print” version of Emotion researcher will now have a more classic look and feel. Though a superficial change, we hope that this adds to the credibility of the articles and other feature pieces published in Emotion Researcher.

Third, we are considering allowing researchers to propose and submit feature articles to be published in Emotion Researcher. In doing so, our aim would be to keep such pieces distinct from what can found in other journals. Additionally, we will need to consider how best to review, edit, and manage such submissions given our limited time and resources. This idea is still being bounced around and we are open to any thoughts or ideas that ISRE members may have.

In the end, Emotion Researcher remains what the ISRE membership makes of it. The upcoming ISRE Meeting will provide an excellent opportunity for us to discuss how to maintain and promote Emotion Researcher as a useful tool and impactful outlet. We look forward to working with all of you to ensure that Emotion Researcher accomplishes these goals.

Sincerely yours,

Carolyn & Eric



Carolyn Price is Senior Lecturer in Philosophy at the Open University (UK). Her research addresses a broad range of questions about emotions – what they are, what they tell us about the world, the norms by which we evaluate them, and (most recently) their relation to the self. She is also interested in particular types of emotions, – such as love, grief and regret. Her book *Emotion (Polity)* appeared in 2015.



Eric Walle is an Assistant Professor of Psychological Sciences at the University of California, Merced. His theoretical research emphasizes the functions of emotions, particularly in interpersonal contexts. Empirically, he examines emotional development, principally in infancy and early childhood, as well as how individuals perceive and respond to emotional communication.

Emotions, One, Two, Many

Arvid Kappas

Department of Psychology
Jacobs University Bremen
a.kappas@jacobs-university.de

As it turns out, empathy is a topic close to my heart and one that has played an important part in my research agenda over the last three decades, so I would like in my last ISRE Matters column to just mention some of my own thoughts and work, how I came to the view that empathy is a construct central to emotion, because the notion of emotion in the individual without at least an implicit social context is basically flawed. I'll then say something about how one can extend concepts that relate to emotion processes between two people to the much more complex situation of how emotions spread on the Internet to form collective emotions; and finally I'll describe a project where we tried to teach empathy to robots so that they can be better tutors for children.

One

My initial views about how to study emotions were very much influenced by Klaus Scherer, who was my mentor and supervisor in Giessen, Germany, when I studied Psychology; and by Paul Ekman with whom Scherer's group had frequently been in contact at the time. Based on the assumption that social context tends to influence the expression of emotions, it appeared best to isolate subjects by placing them in a chamber and presenting stimuli or tasks while they were alone, so that one could see the emotional reaction, apparently untainted by social noise, display rules, or cultural context. In other words – emotions would be studied in individuals cut off from the social context.

Having said that, this was not the approach taken in the experimental paradigm that I used (together with Ursula Hess as the basis of my Masters research (1984-86). Rather, Ursula and I used a mock videophone interaction where the

stimulus person appeared to talk to the subjects one-on-one. We were interested in how changes in the voice, intonation, and facial activity would affect people's perception of emotions and attitudes (see Hess, Kappas, & Scherer, 1988 – for details on that research). We felt at the time that emotions were more likely to happen in social situations and that the communication of these emotions would affect the ongoing interaction, typically depicted by Scherer and his colleagues as a Brunswikian Lens Model.

Two

When I moved to Dartmouth College in 1986 to work with John Lanzetta, I was very much affected by the way in which John, who was an engineer before he was a psychologist, would think of interaction. Everything was about closed loops and feedback processes. On this approach, [empathy was a critical element for the regulation of dyadic interaction](#). Rather than imagining interaction to resemble a ping-pong match, where first Person A communicates with Person B and then Person B with Person A, it appeared better to think of interaction as a closed system with several *concurrent* feedback loops. The idea that social context might affect expressions, and that expressions might impact subjective experience and physiological responses of emotions via facial feedback, led me to propose the [super lens model in 1991](#).



Figure 1. The classic paradigm in experimental psychology to assess affective responses isolates the participant from social influences at a physical level. However, research (e.g., Fridlund, 1991) suggests that implicit social effects remain (photo: Jacobs University).

Because I assumed that we tend to automatically empathize with people, I have developed a view critical of the “classic experimental paradigm” of isolating subjects and confronting them with non-social stimuli. In this context, I consider that the subject, cut off from social interaction in a rather artificial manner will quickly become a [free monadic radical](#) (Kappas, 2013) that is ready to connect to any social context (typically implicit) that is available – for example the experimenters, other waiting subjects, etc. In that sense, and following Fridlund’s suggestion, I assumed that we always display expressive behavior to the people in our head (Fridlund, 1991; see also Hess, Banse, & Kappas, 1995).

These days I believe that empathy is *not* always automatic, but moderated by the social relationship we have to a person. This could be understood as an ingroup/outgroup phenomenon – that is, as depending on whether someone is inside our moral circle or not. A different facet is [how much humanity](#) we grant our explicit or implicit interactant (see Krumhuber et al. 2015). A corollary of the belief that emotion and empathy are always associated with closed feedback loops, is my tenet that it is difficult, and perhaps [not useful, to distinguish between emotion and its regulation](#), as there are always regulatory processes, both social and within the individual, that are at work (Kappas, 2011).

Many

The leap from psychological processes in the individual to dyadic processes is already a great

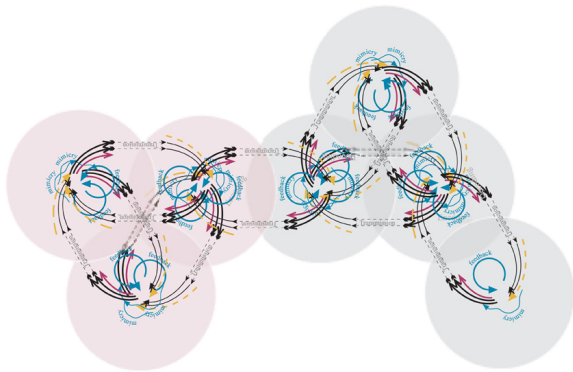


Figure 2. Hyper lens model. Nested communication systems between seven people featuring feedback loops. From Kappas (2013).



Professor Arvid Kappas, May 2017

challenge for the experimental emotion researcher. But increasingly we interact with [large numbers of people over the Internet](#). Not all of these processes are in real time, but they clearly affect us. Research in this area is truly challenging. The [CYBEREMOTIONS](#) project studied how emotions are elicited, communicated, and spread over social networks (Garcia et al., 2016). In this context we can think about [networks of empathic processes](#) – a fascinating new area of research, but one that is particularly challenging due to the new methods that are required to measure and analyze changes in affective states. At a time when political decisions are a function of how many people react to mediated material, such as tweets, images on news sites, or in social media or comments to articles or blogs, it becomes important to understand how and why we seem to emphasize with some but not others; in some circumstances, but not others; and with different time courses. We know way too little about processes at this scale.

Enter the Robot

I, in the meanwhile, have started to teach empathy to robots. Together with a group of excellent researchers from different disciplines and countries, the [EMOTE](#) project developed robotic tutor systems that are designed to respond to the child’s affective state and, for example, adjust teaching strategies as a function thereof. To me, this was a brilliant exercise, one that started from the question of how to define



Figure 3. Arvid explains empathy depicted on the display of a pavilion on Lisbon's Praça do Comércio on the occasion of the ICT 2015.

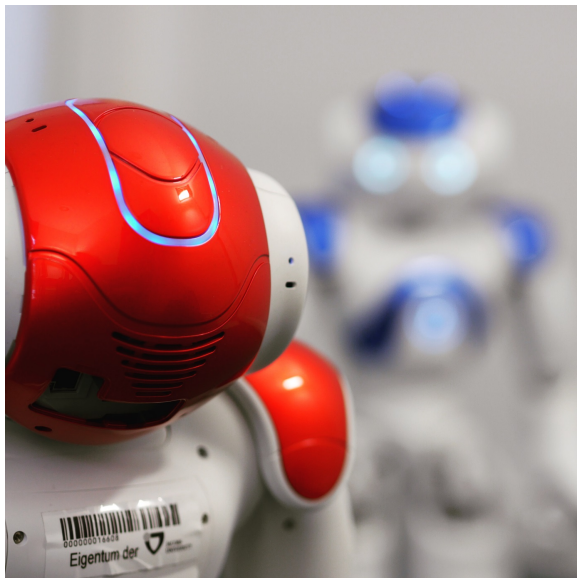


Figure 4. Much of EMOTE's research was using NAO robots (Softbank Robotics) as a platform to study artificial empathic tutors.

empathy in a way that would allow one to build empathic systems, even when there is no commonly agreed upon definition of empathy; and then moved on to investigating how to give the robot advantages over what a human might do (e.g., access to physiological activation data) to compensate for the fact that humans are much better than robots at really understanding a situation, and its implications, as well as visible

affective responses. This research has attracted a lot of attention: we were, for example, invited to the CES in Las Vegas in 2016 to present our work as part of a session on transforming education; and we were featured by the European Commission at its conference on ICT research that it supported.

EMOTE is over, but I have recently received a grant to continue this research in the context of a European Training Network that will fund 15 PhD students distributed over labs from Portugal to Sweden, from 2017 to 2021. If you know of a psychology student who will have a Master's degree at the end of the year 2017 who would want to work in this context and has perhaps some skills relating to robotics – please direct them to me.

[Check out this video on our research](#)

Closing Remarks

This is very likely my last presidential column in the *Emotion Researcher* as the time of my tenure as ISRE's president comes to a close after four years (2 x 2). It has not always been smooth sailing, but always exciting. I am deeply grateful for having had the possibility to serve the society. ISRE was founded while I was a student in Klaus Scherer's lab (see above) and I was immediately impressed by the concept of a society so international, and at the same time so interdisciplinary. At the time membership and attendance at ISRE conferences was restricted to senior researchers and the founding group still reads like a Who's Who. We have opened the society to researchers who have not yet received their PhD and we have made it easier to join. We have started to use social media and creative forms of communicating. When I joined ISRE, the newsletter was a small black and white printed thing of perhaps 8 pages and now its heir – *Emotion Researcher* – is a singing and dancing Internet offering – thanks to Andrea Scarantino's overhaul of the concept. I am thankful for your enormous work, Andrea! This will shape our communication to the rest of the world for years to come. Thank you also for the detailed feedback on my columns over the years. They tend to jump like a bunny from topic to topic and have a very "spoken" style. Thus, my writing always benefitted from your comments. I am very excited that we have two new editors of the

Emotion Researcher: Carolyn Price and Eric Walle. This is the first issue edited by the two. Carolyn and Eric, I wish you all the best. Keep it relevant, keep it up-to-date.

At the St Louis conference from July 26-29, I will thank more people. I hope to see you there. If you have not decided to come – change your mind, we have an excellent conference lined up and we want you there. The next US conference is likely in four years, so now is a good time. It's not too late to book those seats and join us in the Chase Park Plaza Hotel. More info elsewhere in this issue of Emotion Researcher.

References

- Fridlund, A. J. (1991). The sociality of solitary smiles: effects of an implicit audience. *Journal of Personality and Social Psychology*, 60, 229–240.
- Garcia, D., Kappas, A., Kuester, D., & Schweitzer, F. (2016). The dynamics of emotions in online interaction. *Royal Society Open Science*, 3: 160059. ([PDF](#))
- Hess, U., Banse, R., and Kappas, A. (1995). The intensity of facial expression is determined by underlying affective state and social situation. *Journal of Personality and Social Psychology*, 69, 280–288.
- Hess, U., Kappas, A., & Scherer, K. R. (1988). Multichannel communication of emotion: Synthetic signal production. In K. R. Scherer (Ed.), *Facets of emotion: Recent research* (pp. 161-182). Hillsdale, NJ: Erlbaum.
- Kappas, A. (2013). Social regulation of emotion: Messy layers. *Frontiers in Psychology*, 4(51), 1–11. ([PDF](#)).
- Kappas A. (2011). Emotion and regulation are one! *Emotion Review*. 3, 17–25.
- Kappas, A. (1991). The illusion of the neutral observer: on the communication of emotion. *Cahiers de Linguistique Française* 12, 153–168. ([PDF](#))
- Krumhuber, E. G., Swiderska, A., Tsankova, E., Kamble, S. V., & Kappas, A. (2015). Real or artificial? Intergroup biases in mind perception in a cross-cultural perspective. *PLoS ONE*, 10, e0137840. ([PDF](#))

ISRE Biennial Meeting Update

The 2017 ISRE Meeting

Hillary Anger Elfenbein

Conference Host
Olin School of Business
Washington University in St. Louis
helfenbein@wustl.edu

The biennial conference of the International Society for Research on Emotion will be held in St. Louis, USA from July 26th through 29th. ISRE brings together researchers from all the disciplines involved in the study of emotion from all over the world to meet and share their work.

St. Louis is a thriving metropolitan area filled with cultural riches to explore. The city was a transportation hub in the days of the railroad and Mississippi riverboats, and was considered the “Gateway to the West” as the starting point for wagons on the Oregon Trail. The conference will be held in the city’s Central West End, which has a European feel with small walkable streets and large sidewalks full of outdoor dining. Eateries feature upscale, local, and ethnic cuisine. Our venue will be the famed Chase Park Plaza, which often ranks as one of the best hotels in the US due to its luxurious renovation that preserves historical details of the landmark property. Washington University in St. Louis is a top school in the US and a generous supporter of the ISRE conference. St. Louis has easy flight connections across the US and to major airline hubs overseas.

Looking forward to seeing you there!



Historic Chase Park Plaza, site of the 2017 ISRE Meeting.



The surrounding West End Neighborhood.



The World Chess Hall of Fame, located in the West End Neighborhood.



Forest Park is across the street, 1300+ acres with trails, a zoo, and an art museum.

ISRE Biennial Meeting Update

The 2017 ISRE Meeting Program

W. Gerrod Parrott

Program Committee Chair
Department of Psychology
Georgetown University
parrottg@georgetown.edu

On behalf of the Program Committee, I am pleased to welcome emotion researchers to the upcoming meeting of ISRE, which will convene in the USA in the city of St. Louis, Missouri on 26-29 July 2017. The Program Committee received a strong set of submissions which have been arranged into plenary sessions, symposia, panels of talks organized around themes, poster sessions, and a session of rapid-fire presentations of late-breaking research.

We are pleased to have achieved several noteworthy goals in setting up this year's program. The various sessions represent the full range of disciplines now engaged in research on emotion, and this year's conference will be an opportunity for attendees to learn about recent research from a variety of perspectives. There are speakers from philosophy, sociology, history, literature, computer science, linguistics, and education, as well as from the full range of psychological approaches – social, cognitive, developmental, neuroscience, cultural, industrial/organizational, and clinical. The Program Committee achieved this goal without issuing special invitations or mandating a pre-set set of topics; this range of approaches and topics emerged naturally from the papers that were freely submitted by the diverse ISRE membership, and thus represents a true sample of the approaches and topics that are active and of interest this year.

We have invited three plenary speakers and a special Presidential Symposium, one of which is programmed for each day of the conference. Jonathan Gratch, a computer scientist and psychologist, will speak at the opening session on the topic, "The promise and peril of emotionally intelligent machines." The opening speaker on

the first full day of the conference will be Ruth Leys, a humanities theorist. Her talk is titled "Outside-in: Mirror neurons and the social performance of empathy." The speaker at the plenary session on the second full day will be Lynn Smith-Lovin, a sociologist who will speak on "Identities, selves and mixed emotions." On the final day of the conference, ISRE President Arvid Kappas will chair a symposium on the role of empirical evidence (and replication thereof) for theories of psychology; the panelists will represent the fields of philosophy, psychology, and sociology. To mention just a sample from the many fascinating topics on this year's program, there will be presentations about facial expressions and emotional communication, political and moral emotions, depression, appraisal, affective computing, history of emotions, emotional development, social exclusion, and a host of specific emotions.

We are especially pleased to have involved researchers who are early in their careers. Special discount rates for registration and hotel accommodation have made attendance more economical. And the variety of presentation formats – talks, posters, rapid-fire presentations – have provided an appealing range of options.

The program is online. Check it out at <http://isre2017.org/program.html>.

The variety of interesting topics is most exciting. I'm pleased that we were able to arrange the sessions so that there are never more than four sessions occurring simultaneously, and usually there are only three. Nevertheless, if you are like me your problem is going to be in deciding which session to attend and which regrettably to miss.

I look forward to seeing everyone in St. Louis!



In Memoriam

Personal Reflections on the Neuroscientific Legacy of Jaak Panksepp (1943-2017)

Douglas F. Watt

School of Medicine
Boston University

“Nature has placed mankind under the government of two sovereign masters, pain and pleasure . . . they govern us in all we do, in all we say, in all we think: every effort we can make to throw off our subjection will serve but to demonstrate and confirm it.”

Jeremy Bentham

One of the truly great voices in modern neuroscience was silenced recently when Jaak Panksepp suddenly died on April 18th of this year, following a fairly routine surgical procedure that went badly awry, generating a serious ischemic-hypoxic insult, a tragedy from which not even Jaak Panksepp could recover. Jaak had battled long and hard in relationship to recurring bouts with various cancers, and is too often the case, it was actually the cancer therapy process that eventually killed him and not the cancer itself, although that distinction offers little solace. Nevertheless, the terrible and final suddenness of all this, despite Jaak’s age and medical history, was shocking and deeply upsetting to all of us who were close to Jaak. None of us is truly prepared for these departures at the end of life, even though we all know that they are coming eventually... and inevitably. Jaak’s neuroscientific work was, as many appreciate, centered at least in part around separation distress as a prototype emotional state, and he has had much to say, over the many years of his career, about the nature of the very grief that all who knew Jaak well have been feeling since his passing. That grief still lies heavy on many hearts, of family, friends and colleagues.

The personal and professional losses for me are almost impossible to separate, as I lost a senior colleague, a highly valued friend, and a mentor in neuroscience, who was decisively the largest and most positive single influence in the second half of my professional career. But neuroscience lost at least as much, an irreplaceable leader, a brilliant scholar and scientist, and a most passionate advocate for an affect-centered view of the brain and mind. I take solace in my sense of great good fortune for having had 20+ years of close colleagueship and friendship with Jaak, sharing half a dozen week-long Affective Neuroscience seminars, many book chapters, dozens of larger conferences and other fruitful and sundry projects. I believe that over time his legacy and many gifts to the sub-fields of mind and brain science will prove increasingly resonant and prescient. Progress in science is intrinsically slow by its very nature, but once in a while a highly creative genius comes along, whose insight and integration of large amounts of previously fragmented material, creating new paradigms and concepts, moves the field forward in sudden and deceptive leaps. In the increasingly ‘twinned’ fields of neuroscience and psychology, Jaak was one of those people.

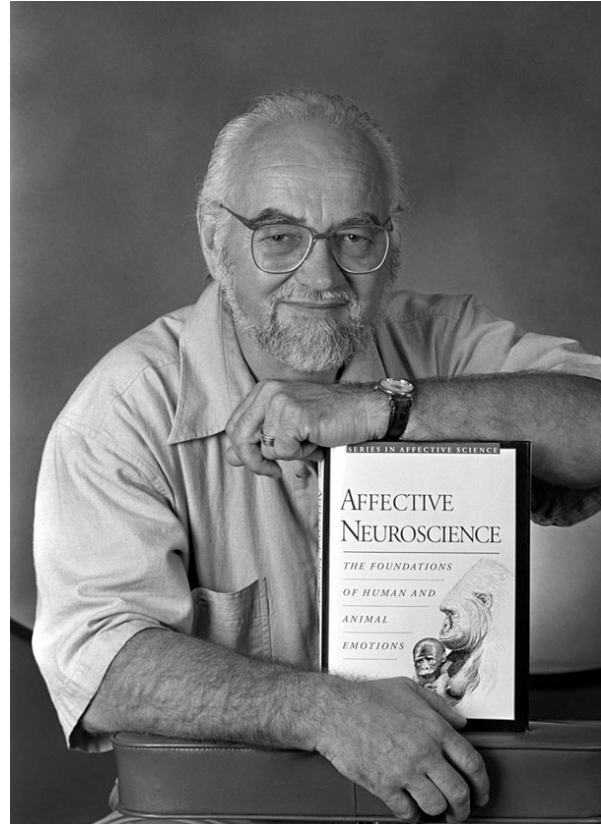
By any conventional standard of scientific productivity, Dr. Panksepp was a giant contributor in relationship to the neuroscience of emotion and clearly staggeringly prolific, with over 450 empirical and review publications, and literally many dozens of book chapters in various edited volumes and compilations. His publication vitae centrally featured his 1998 textbook, *Affective Neuroscience*, increasingly regarded as a landmark and a classic publication in neuroscience. This was updated with Lucy Biven in the less technical and more accessible treatment, *The Archaeology of Mind*. He also solo edited the well-regarded two-volume *Advances in Biological Psychiatry* in 1996, the more recent *Textbook of Biological Psychiatry* in 2004, and the highly technical compilation *Handbook of the Hypothalamus* with Peter Morgane which few have read but which, despite its age, is still to this day the most comprehensive collection of work on that critically important neural system and its functional networks. The very last of his many edited volumes, *The Psychology and Neurobiology of Empathy*, was

completed with me in 2016. This huge body of work has had an enduring and pervasive influence on the fields of psychology, psychiatry, and both clinical as well as experimental neuroscience, and has inspired at least two generations of young scientists eager to understand more about the mysteries of the brain and mind and the pivotal role of emotion in its neurodynamics.

Where and what were Jaak Panksepp's greatest contributions? This is not an easy question to answer, because his interests spanned across many traditional disciplinary boundaries, including such disparate topics as a theory of music, the fine-grained anatomy and connectivity of the mesodiencephalon, and the nature of laughter as an affective signal. And yet the common denominator in all of these was Jaak's keen and unwavering focus on the *nature of affect and the implications of affect for every other important psychological and neuroscientific territory*. The following list of Jaak's major areas of scientific contribution has to be considered simply my personal scientific opinion, and therefore hardly definitive or unbiased, but I will outline three main areas of scientific contribution. Others, indeed, many others might be added, but space considerations demand a shorter list – so here's mine.

On a fundamental and intrinsic relationship between consciousness and emotion

As perhaps his most critical contribution, Jaak continually emphasized the centrality of emotion and affect in any potential neuro-architecture for a conscious mind, what I would call an '*affect-centric*' view of the embodied mind. *Affects were evolutionarily conserved core routines guiding the living mind*. This strong conviction that emotion provided not just an interesting 'coloration' to consciousness contrasted starkly with the cognition-centered and the predominantly sensory view of consciousness at the time of his early seminal work and publications. Instead of this conventional sensory/cognitive perspective on conscious architectures, *affects were conceptualized as signaling functional integrations foundational to any version of a*



Dr. Jaak Panksepp (Photograph courtesy of Bowling Green University)

conscious architecture; additionally, all strong affects were thus intrinsically conscious, a shared scientific conviction that originally brought us together and created an immediate intellectual bond that was to energize our many shared ventures and projects.

Thus it was no coincidence that we met for the first time in the second biannual Tucson Series on the Science of Consciousness in 1996. It would be an understatement to suggest that affective neuroscience perspectives might have been underrepresented at the Tucson Conference in 1996. We shared an immediate sense of dismay, tempered perhaps by a certain amusement, that the fields of psychology and neuroscience seemed at times so blissfully unaware that *a mind without the internal 'value compass' provided by emotion and affect had no basis for organizing basic behavioral, or for that matter, higher attentional and cognitive priorities*. "What are they smoking?" we wondered over drinks at that first conference!

We shared a sense of astonishment that affects were not more universally appreciated as the true ‘reinforcers’ referenced by behaviorism, and wondered how anyone in the many disparate fields of psychology and neuroscience could somehow miss this central ‘keystone’ fact of our psychology. As consciousness had been for the four previous decades a mostly discredited subject under the still powerful aegis of behaviorism, we were joined in this early conviction in the mid-1990s by precious few in neuroscience, but notably, by Antonio Damasio. Damasio, Joe LeDoux, and several other luminaries in emotion theory and consciousness studies all participated as commentators in one of my first ventures with Jaak, an ASSC (Association for the Scientific Study of Consciousness) web-seminar on Emotion and Consciousness given in the summer of 1998, which generated lively debates still resonating to this day about exactly how to conceptualize the role of emotion in consciousness. Can consciousness and emotion be simply conceived of as largely orthogonal processes, with just a single intersection, when emotion created a conscious affective state? We argued that this single intersection was just the tip of the iceberg, in terms of the rich interweaving of affective signals into every aspect of conscious life, from what we found interesting, to what we avoided, to what we craved and sought out, from what we learned forever to what we forgot next week, indeed, virtually every aspect of the mind.

The crux of our early shared conviction was simply that affects (both homeostatic and emotional) and their underlying brain systems, *had to be foundational in some poorly understood fashion for how neural activity gives rise to an organized agency* – a ‘selfhood’ in the brain, and furthermore, that *such a self would be an ‘agentic self’ and not any version of a passive observer, with behavioral initiatives and basic motivation intrinsically organized around and by those emotional and homeostatic affects*. In other words, it made sense to assume that all behavior was forever hinged to the pursuit of the many and varied positives and also equally in the service of minimizing the many varied negatives, echoing Bentham’s early brilliant intuition, an insight essentially recapitulated by Freud 100 years later.

Without those ‘value signals’, there was simply no other basis for motivation or behavior.

This set of early assumptions – which ran largely counter to the sensory and cognition-centric views in early consciousness studies – is being slowly but progressively validated, but we found an unlikely early ally in the global workspace perspectives of cognitive theorist Bernie Baars and colleagues, such as Jim Newman. This central neurologic dependency of conscious architectures upon affective signals was revealed in the neurodynamic dependence of cortical systems on the integrations achieved in the mesodiencephalon to do any form of meaningful and organized cognitive work (see Merker, 2007 for the most definitive and exhaustive summary to date). It was also revealed starkly in the underappreciated syndrome of akinetic mutism, where collapse of emotion and motivation leaves a virtually rudderless and engine-less thalamocortical system and a disorder of consciousness (see Watt and Pincus, 2004 and Watt, 2012 for reviews). And while one can say that neuroscience still lacks its ‘keystone in the arch’ – a definitive and well-validated neural theory of mind as one of the core emergent properties and mysteries in Nature – increasing evidence argues that conscious architectures are highly distributed and non-local. They exist as transient adaptive networks, flitting about in an ‘enchanted loom’ in Sherrington’s haunting metaphor, with new transient and highly distributed alliances being formed on a moment-to-moment basis, but bridging the entire vertical axis of the brain.

At this point in the evolution of a neuroscientific theory of mind, the critical role of the mesodiencephalon in such distributed networks, as upper brainstem and medial diencephalic territories rich in ancient homeostatic and affective systems, is simply beyond meaningful dispute. An extensive lesion of these mesodiencephalic territories will collapse consciousness totally and permanently, even without a trace of damage to any aspect of the more telencephalic thalamocortical systems that are virtually the exclusive anatomical focus of cognitive neuroscience. Indeed, just a full lesion of the tiny territory of midbrain PAG – at the ventral tip of the large group of brain affective systems running from paleocortex through limbic

regions, to hypothalamus and down into the brainstem, leaves the person ‘gone’ in a comprehensive sense, and with a permanent and severe akinetic mutism. This may hint at PAG’s role in laying the neuroevolutionary foundations for those more evolved affective systems, sitting at the bottom of the highly distributed affective networks extending upward from reticular-hypothalamic to limbic and paralimbic systems.

The neurodynamics of this profound ‘system integration’, particularly in terms of a developmental trajectory that bootstraps the conscious mind from core homeostatic and affective systems, clearly remain to be more fully elucidated, but the ontogenetic development of consciousness for each of us appears to clearly recapitulate our phylogeny, as the brain builds networks from the inside (ventral and medial regions) out to the more telencephalic lateral and dorsal regions. Although the cortex still dominates cognitive neuroscience – and with good reason as it forms the neural substrate for virtually all cognitive content – I believe that future generations will strongly credit those handful of neuroscientific pioneers who kept pointing at neurodynamic integrations, still poorly understood, but achieved first in more ancient mesodiencephalic, tectal and tegmental systems, as *foundational for all the cognitive activity at the top of the system*. Jaak will likely be highly regarded and credited as one of those prescient pioneers.

On understanding emotion in terms of interactive but partially discrete prototype states energized by an overarching, precedent and integrative SEEKING system

Jaak Panksepp also made pivotal contributions to unraveling one of the central mysteries of emotion, namely, that it somehow functions as *a central clearinghouse in which all emotional activators must compete and interact*, a fact which necessitates global and highly distributed neural system architectures for emotion (and an underappreciated direct correlate of the global network ideas just reviewed). In other words, emotion cannot be allocated to any kind of simple ‘module’ or any version of a local network, but has to be understood as a distributed

integrative system, a realization which made him rather fond of placing the affective primes in all capitals (‘FEAR’, ‘RAGE’, ‘GRIEF’, ‘PLAY’, ‘SEEKING’, ‘CARE’), as his way of denoting this distributed or ‘system’ property. Although this might have seemed semantically ‘quirky’ to some (including even to this author at times ☺), it was energized by Jaak’s realization that more global system concepts were mandated, and where perhaps the various affective primes might function as ‘nodes’ within an integrated system, thus creating a ‘neural clearinghouse’ in which all the potential affective states could compete and interact in a variety of adaptively critical antagonistic and agonistic ways. And Jaak would truly cringe when someone might extrapolate that each of these capitalized primes might be some kind of ‘module’ in the brain. In his view, the negatives and positives would have to be understood as fundamentally antagonistic, but there were also more subtle versions of promotion and inhibition between systems of similar valence. For example, while playfulness might ‘set the table’ so to speak for the activation of sexual desires, at some point excessive playfulness might actually inhibit activation of the LUST system. Such an image of complex agonistic and antagonistic systemic interaction between the primes fits beautifully with the phenomenology of emotion and emotional behaviors, as all kinds of emotional pushes and pulls clearly do agonize and antagonize one another in various ways – part of what gives our (and general mammalian) behavior its delightful and frustrating unpredictability.

In this sense, I believe many, particularly the social constructionist school of emotion theorists, did not fully appreciate how in Panksepp’s system of the seven emotional primes, the SEEKING system was conceptualized as a “special class of one”, and that it had to be the evolutionarily precedent or ‘master’ system, a supposition recently gaining impressive validation by empirical work on dopamine in crayfish (Huber et al., 2011). In social constructional theories of emotion, the notion of ‘core affect’ (i.e., Russell and Barrett, 1999), with an emphasis on dimensional aspects of approach/avoidance, arousal, and valence, is not actually operationally dissimilar to how Panksepp conceived of the integrative and

Affective Prototype	Distributed Neural Networks and Major Structures	Neuromodulators
Generalized Motivational Arousal – SEEKING	Ventral Tegmental Area (VTA) to lateral hypothalamic to periaqueductal gray (PAG), with diffuse mesolimbic and mesocortical “extensions.” Nucleus accumbens as crucial basal ganglia processor for emotional “habit” systems and affective learning.	DA (+), glutamate (+), many neuropeptides including opioids, neurotensin, CCK, and many other facilitators
RAGE (Affective Attack)	medial amygdala to bed nucleus of stria terminalis (BNST) to anterior and ventromedial and perifornical hypothalamic to more dorsal PAG	Substance P (+) (? Ach, glutamate (+) as nonspecific modulators?)
FEAR	central & lateral amygdala to medial hypothalamic to dorsal PAG to nucleus reticularis pontine caudalis	Glutamate (+) and neuropeptides (DBI, CRF, CCK, alpha MSH, NPY)
LUST (Sexuality)	BNST and corticomedial amygdala to preoptic and ventromedial hypothalamus to lateral ventral PAG	Sex Steroids (+) (T/E), vasopressin, oxytocin
Nurturance/maternal CARE	Anterior cingulate to bed nucleus of stria terminalis (BNST) to preoptic hypothalamic to VTA to more ventral PAG	Oxytocin (+), prolactin (+), dopamine, opioids
Separation Distress/PANIC (Social Bonding)	Anterior cingulate/anterior thalamus to BNST/ventral septum to midline & dorsomedial thalamus to dorsal preoptic hypothalamic to more dorsal PAG (close to circuits for physical pain)	Opioids (-/+) oxytocin (-/+), prolactin (-/+) CRF (+) for separation distress, ACh (-)
PLAY/ (Social Joy & Affection)	Parafascicular/centromedian thalamus, dorsomedial thalamus, posterior thalamus, projecting to ventral PAG (septum inhibitory re: play)	Opioids (+ in mod. amounts, – in large amounts), ACh (+), cannabinoids (+)

superordinate role of the SEEKING system, although there are still important differences between Panksepp and social construction theories that cannot be glossed over. Panksepp parted company with constructional views if they went so far as to argue that the categorical or ‘emotional prime’ systems (observing for example a principled distinction between PLAY and RAGE as neurobiologically meaningful) were simply a specious cultural convention or ‘meme’ carried over from a pre-scientific ‘folk’ nomenclature, but without any real neurologic and biological basis.

Panksepp’s theory of the emotional ‘primes’ provides real ways of immediately and directly linking the SEEKING system to all the other primes. In other words, FEAR was the seeking of safety, PLAY the seeking of rough-and-tumble joyful engagement, RAGE the seeking of an end (sometimes ‘with prejudice’) to that toxic agent or circumstance that might be the source of one’s frustration or injury, maternal CARE the seeking

of safety, restored homeostasis or good feelings in a dependent and vulnerable other (and thus indexing a proto-empathy system, as classically outlined in Panksepp, 1998, and more recently in Watt and Panksepp, 2016). In this critical sense, I sometimes felt that students as well as some critics of Panksepp did not always clearly appreciate that all of these classic affective ‘primes’ or prototypes had to be, at least in some sense, *specialized ‘resonances’ or evolved ‘nodes’ interacting within a more ancient and precedent generalized motivational arousal or SEEKING system*. Perhaps this key insight was expressed too implicitly at times, and where this precedent role of the SEEKING system was not made more explicit, although it is stated at several points in his magnum opus (Panksepp, 1998).

One underappreciated benefit – indeed one great beauty of this set of concepts from the standpoint of scientific theory – is that it provided an integrated neural substrate for the competitive interaction of the emotional primes, where the

negatives and positives could easily inhibit each other, a clear functional fact of affective life that any good model had to explain. Additionally, the SEEKING system concept provided an easy evolutionary bridge from homeostasis to emotion, as the SEEKING system network contains, as one of its core nuclei, the lateral hypothalamus which receives abundant dopaminergic enervation from the midbrain, providing an immediate neural bridge for integrating behavior (via its extensive basal ganglia and mesolimbic connections), homeostatic imbalances, and sources of homeostatic relief. Thus, in its most ancient forms, as Jaak famously offered, “the SEEKING system gets thirsty animals to water, hungry animals to food, cold animals to warmer environments,” etc. The SEEKING system creates neural links from rewards and punishments to reward predictors or reward cues, and binds both of those classes to basic behavioral routines, such that a hungry animal can activate motor routines to approach a source of food and then initiate consummatory routines, under the guiding influence of this master motivational system. It is hard to imagine a more elegant set of integrative predictive tools about emotion and motivation than what this set of basic ideas affords modern neuroscience.

Another underappreciated benefit of Jaak’s emotional system concepts was that they provided a highly plausible theoretical bridge for how homeostasis proper (likely the antecedent evolutionary process) might have given rise to the emotional primes, as they emerged from basic evolutionary-predictive extensions of core homeostatic mandates. This was no coincidence, as Jaak cut his teeth on energy balance research. In other words, perhaps *fear reflects an anticipatory forward-looking extension of pain and the basic aversion to tissue damage*. In fear, we are not damaged yet by an approaching predator or more powerful rival, but we might be, if we do not seek safety, and either freeze to avoid detection, or flee, if this less metabolically costly solution is not possible. In separation distress, a young infant mammal is not yet metabolically compromised but suddenly finds itself separated from both its source of metabolic supply (caretakers) as well as its primary protection from predation, suggesting that *separation distress*

may have emerged as an anticipatory forward-looking extension of these basic homeostatic routines to ensure the safety of vulnerable offspring. While this way of linking more ancient homeostatic processes with classic emotional primes remains still mostly untested, and in need of much further probing and investigation, there is no meaningful evidence (that I am aware of anyway) against it. It is also deeply consistent with how evolution functions to extend the brain as an increasingly sophisticated prediction engine, stacking new predictions on top of older successful behavioral routines. Evolution operates from the highly conservative principle of having no need to “re-invent the wheel” in every new context, just tweak and enhance the existing operations a bit here and there to manage the new adaptive context. Encephalization is, in this real sense, nothing more than the creation of more and more sophisticated prediction engines (and consistent for example with evidence that Wernicke’s area is specialized for the prediction of phonemic sequences). Jaak’s early work on the mysteries of feeding and energy balance exposed him at the beginning of his career to the challenge of finding integration across highly varied and different classes of motivational ‘vectors’ including evolutionarily more ancient and more recent classes of motivations, and his SEEKING system concepts were the brilliant solution.

This basic set of theoretical notions finds strong validation in the fact that opioids not only modulate pain and reward satiety (two very ancient homeostatic operations presumably evolutionarily preceding the classic emotional primes), but in more complex and extended mammalian brains they have come to modulate social bonding, play and separation distress (all critical emotional operations of a highly social brain). Play and quieter forms of social comfort are high opioidergic states, while separation distress is organized in good part by the inverse condition of low opioidergic signals (validated in humans by Zubieta et al., 2003), coupled perhaps with high CRF signals. In this sense, separation distress might reflect an evolutionary extension of pain, while social comfort in a secure attachment perhaps is an evolutionary extension of a basic homeostatic wellness and satiety. *Thus, the SEEKING system concept provides a ready*

and plausible evolutionary and functional bridge between homeostasis and emotion that no other neural system concept currently affords to my knowledge.

Panksepp was able to successfully interpret the confusing results of the electrical brain stimulation (EBS) experiments on the ‘self-stimulation system’ (as originally labeled by Milner and Olds), and was not seduced by the simplistic behaviorist concept that this was “**the** reward system.” Jaak realized from many experiments that this system had to be conceptualized differently from these seductive oversimplifications, as the data suggested that this system supported *both the seeking of rewards but also the avoidance of punishments*. This SEEKING system concept allowed Jaak to achieve a theoretical integration and coherence where there had been previously only confusion and apparent contradiction, particularly around the way in which electrical brain stimulation of SEEKING system trajectories (such as VTA, accumbens or medial forebrain bundle) could generate such apparently disparate behavioral results. *Jaak realized that the behavioral and affective trajectories off of EBS **should** be variable, given the SEEKING system’s precedent and overarching role*, depending on many variables, including species and animal personality phenotypes, adaptive context and even the current ‘mind set’ or mood of the animal, which would all influence which affective ‘prime’ might become emergent from EBS of the mesolimbic mesocortical DA system. Thus, predatory animals like cats might be much more likely to go into a stalking mode, while forager species like rats might start sniffing around and exploring the environment, while restrained animals might actually just try to escape more vigorously. Unfortunately, this behavioral phenotypic variability resulting from EBS “self stimulation” has been misconstrued by constructional theorists of emotion as hard evidence against the ‘categorical’ or ‘affective prime’ view. But it’s not; it is simply more evidence that not everyone has clearly understood the full implications of Jaak’s complex ‘affective prime’ theory and the precedent, superordinate role of the SEEKING system.

Then there were the many discrete contributions (too many to specifically enumerate

here) Jaak made to understanding each of these affective prototype system, aside from the overarching role of a SEEKING system, most particularly empirical contributions in relationship to separation distress/GRIEF, maternal CARE, the organism ‘defense states’ of FEAR and RAGE and perhaps his favorite system, PLAY. The table below, first published in the ASSC Web-seminar on emotion and consciousness (from Watt, 1998) but then was featured in many of Jaak’s reviews. This table summarizes many of his basic empirical findings (table extracted largely from Panksepp, 1998).

On the likely centrality of mammalian play in the construction of a social brain

Jaak Panksepp’s third area of critical contribution might be around placing the PLAY system more clearly front and center as a critical pro-social system in the mammalian and human brain. There are several indices suggesting that we are perhaps the most socially dependent creature on the planet, with this a hidden corollary to the developmental costs of a big cortex and the attendant and intrinsic helplessness of human infancy. The importance of play and its critical role in a variety of prosocial operations of the human mind and brain is still widely underappreciated. For all the contributions of various branches of psychoanalysis to an affect-centric view of the mind, even psychoanalysis (with some notable exceptions) neglected this prototype emotional system, and sometimes treated play as simply a proxy for aggression and/or dominance under the aegis of its original ‘dual instinct’ theory. There is now so much accumulating evidence that this system is critical in a profound way for humans – it promotes attachment, signals trustworthiness (we trust playful people far more readily and much more immediately than dour, irritable or guarded people), and interacts powerfully in a supportive and facilitating way with LUST and CARE (proto-empathy) systems. Its cognitive extensions in humor ripple through much of our social discourse. We can’t seem to get enough of it in fact. Despite all the endless mass media and pop psychology chatter and confusion around what might create the basis for a good relationship, *both men and women when surveyed about this question want first and foremost*

someone who is playful and has a good sense of humor (Chick et al., 2012). And if the PLAY system in the brain is responsible for all forms of laughter, and if spontaneous smiling (as opposed to ‘social smiling’) is part of the PLAY affect continuum, and ‘on the way’ to laughter so to speak (amplification and strengthening of the ‘pro-smiling’ stimulus leads to outright laughter), *this might well be the most underappreciated pro-social neural system in the human brain*. If spontaneous smiling and laughter are essential social signals which indicate that “things are indeed okay,” if playful interactions are a powerful augments of attachment, a predictor of who we are going to like and want to be with, and also potentially inhibit the stress axis of the brain, Jaak’s advocacy for the critical importance of this highly prosocial system, at a time when almost no one else in neuroscience was paying much systematic attention to play, playfulness, or humor, looks more and more prescient.

What might my overall conclusion be from this body of work? Certainly many different individuals might generate many different take-home messages from Panksepp’s work, but for me, as I stated at the beginning, it is virtually impossible for me to separate the personal from the professional in relationship to Jaak and his work. The most enduring aspect of Jaak’s legacy for me simply is something rarely made explicit in his work, but which we talked about on a number of occasions as being more implicit – a deep reverence for an animate Nature, which may both energize the best science and may be an emotional endpoint of the best science as well. Spirit appears to be matter profoundly organized into truly dizzying levels of complexity – that such wondrous properties as minds with affects emerge from a physical Nature over a long and still in many ways mysterious evolutionary course deepens a sense of awe and wonder. As Panksepp himself stated many times, affects are deep and profound evolutionary gifts, ‘voices of the genes’ in his phrase, conserved value signals that provide our only real compasses keeping us on track in life, if we can understand, regulate and integrate them. Although I feel a keen sense of loss at his passing, I am also very grateful for having encountered Jaak Panksepp in my life’s serendipitous travels, and for the many personal

and intellectual treasures that have emerged from that relationship.

References

- Chick, G., Yarnal, C., & Purrington, A. (2012). Play and Mate Preference: Testing the Signal Theory of Adult Playfulness. *American Journal of Play*, 4, 407–40.
- Huber, R., Panksepp, J. B., Nathaniel, T., Alcaro, A., & Panksepp, J. (2011). Drug-sensitive reward in crayfish: An invertebrate model system for the study of SEEKING, reward, addiction, and withdrawal. *Neuroscience & Biobehavioral Reviews*, 35, 1847-1853.
- Merker, B. (2007). Consciousness without a cerebral cortex: A challenge for neuroscience and medicine. Target article, commentaries and author’s response. *The Behavioral and Brain Sciences*, 30, 63-134.
- Morgane, J. P., & Panksepp, J. (1981). *Handbook of the Hypothalamus: Vol. 4: Part B. Behavioral Studies of the Hypothalamus*. New York: Marcel Dekker, Inc.
- Morgane, J. P., & Panksepp, J. (1980). *Handbook of the Hypothalamus: Vol. 3: Part A. Behavioral Studies of the Hypothalamus*. New York: Marcel Dekker, Inc.
- Morgane, J. P., & Panksepp, J. (1980). *Handbook of the Hypothalamus: Vol. 2: Physiology of the Hypothalamus*. New York, NY: Marcel Dekker, Inc.
- Morgane, J. P., & Panksepp, J. (1979). *Handbook of the Hypothalamus: Vol. 1: Anatomy of the Hypothalamus*. New York, NY: Marcel Dekker, Inc.
- Panksepp, J. (2004). *A Textbook of Biological Psychiatry*. New York, NY: Wiley Press.
- Panksepp, J., & Biven, L. (2012). *The Archaeology of Mind: Neuroevolutionary Origins of Human Emotion*. New York, NY: W. W. Norton & Company.
- Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. New York, NY: Oxford University Press.
- Panksepp, J. (1996). *Advances in Biological Psychiatry, Vol. 2*. Greenwich, CT: JAI Press.
- Panksepp, J. (1995). *Advances in Biological Psychiatry, Vol. 1*. Greenwich, CT: JAI Press.

- Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, 76, 805–819.
- Watt, D. F. (2012). Theoretical challenges in the conceptualization of motivation in neuroscience: Implications for the bridging of neuroscience and psychoanalysis. In A. Fotopoulou, D. W. Pfaff, & M. A. Conway (Eds.), *From the Couch to the Lab: Trends in Psychoanalysis* (pp. 85-108). New York, NY: Oxford University Press.
- Watt, D. F. (1998) Emotion and consciousness: Implications of affective neuroscience for ERTAS theories of consciousness. Internet Electronic Seminar for August–September 1998. *Association for the Scientific Study of Consciousness*.
<http://server.phil.vt.edu/assc/esem.html>.
Available from author by request.
- Watt, D. F., & Panksepp, J. (2016). *The Neurobiology and Psychology of Empathy*. New York, NY: Nova Science Publications.
- Watt, D. F., & Pincus, D. (2004). Neural substrates of consciousness: Implications for clinical psychiatry. In J. Panksepp (Ed.), *Textbook of Biological Psychiatry* (pp. 75-110). New York, NY: Wiley Press.
- Zubieta, J. K., Ketter, T. A., Bueller, J. A., Xu, Y., Kilbourn, M. R., Young, E. A., et al. (2003). Regulation of human affective responses by anterior cingulate and limbic mu-opioid neurotransmission. *Archives of General Psychiatry*, 60, 1145–1153.

ISRE Interview

The Social Construction of Emotion: Myths and Realities

James R. Averill

An interview with Andrea Scarantino
(July 2017)

James (Jim) Averill is Professor Emeritus in the Psychology Department at the University of Massachusetts, Amherst, and past President of the American Psychological Association's Division of Theoretical and Philosophical Psychology. He has written more than 130 articles, book chapters and reviews on various aspects of emotions, most influentially on physiological markers of emotions, the language of emotion, and emotional creativity, as well as on specific conditions as grief, stress, anxiety, love, anger, hope, happiness, solitude, and aesthetic experiences. Averill is widely considered one of the founding fathers of the social-constructionist approach to the study of emotions, which characterizes emotions as transitional social roles and questions the assumption that they just unwittingly "happen" to us.

Where did you grow up? What did your parents do? How was your family like? Do you remember what your career dreams were as a young man?

Until I was 9 years old, I grew up in the foothills of the Sierra Nevada mountains, east of San Francisco. My parents ran a winter and summer resort called Pinecrest, on a lake by the same name. In the summer, we stayed at the resort and in the winter in Sonora, the largest nearby town. The resort was built by my mother and her brother, but he died young. My father met my mother while working for the government, surveying public lands (e.g., national forests). They had three children, two daughters and myself, the youngest. When I was about 8, I came down with bulbar polio. It was not diagnosed correctly at the time: there was no obvious paralysis, except that I had difficulty swallowing



Jim Averill, 2017

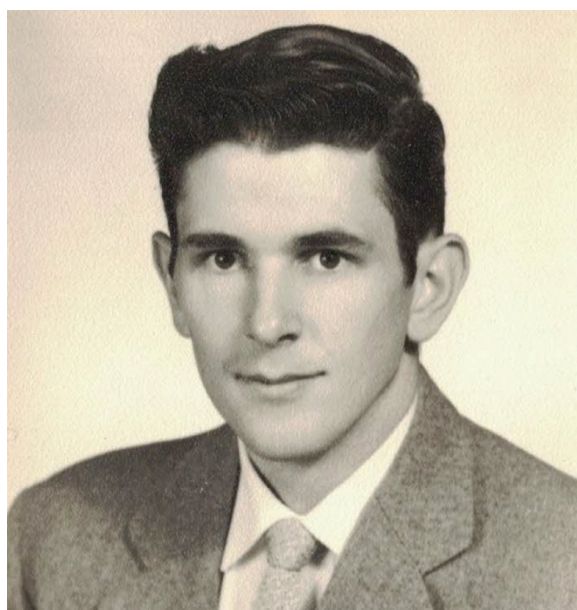
and periodically lost consciousness. The assumption was that I had encephalitis. My eldest sister came down next with spinal polio, which left one leg paralyzed. Diagnosis was then obvious. My parents were advised to move to a warmer climate for my sister's sake. We therefore moved to Glendale, a suburb of Los Angeles, where my father resumed work as government surveyor; he was in the field, away from home, about six months out of the year. My mother therefore managed the household. When my father retired, we moved to Oceanside, north of San Diego, where my father had grown up.

As far as my career dreams as a young man are concerned, they were never coherent. It took a considerable time and much experimenting before settling on a career. After graduating from High School, I spent the summer quarter attending Mexico City College. Instruction was in English, so I learned little Spanish. In general, I am very poor at learning languages. However, I did gain an appreciation for Mexico: It is a country with rich potential, yet to be fully realized. Returning to Oceanside, I didn't know what I wanted to pursue as a career. I therefore enrolled in the local community college. At the end of a year, I had taken all the courses that interested me, yet I still couldn't decide on what to do next. So, in 1954, I joined the Army. It was

the tail-end of the Korean war, but the armistice had not yet been signed. By joining when I did, I not only preempted the draft, but became eligible for the Korean G.I. Bill, which would help pay for college when I got out. The Army sent me to language school for a six-month immersion course in German, which is somewhat ironic considering my poor aptitude for learning languages. I spent the remainder of my tour stationed in Frankfurt, Germany, where I had ample opportunity to travel throughout much of Western Europe.

In 1959-1960, you spent one year at Düsseldorf Medical Academy and University of Bonn, Germany. Why did you go to Germany, and what are your memories of that experience abroad?

When I was discharged from the Army, I still had no clear career objective. Procrastinating yet again, I enrolled at San Jose State College (now University), which had a combined program in philosophy and psychology that looked interesting. I enjoyed both disciplines, so I split the major and got degrees in both psychology and philosophy. Now, I had only to decide between them as a career. I doubted that I could make a meaningful contribution to philosophy, a field with a rich history stretching back several thousand years. I therefore applied to the



To Germany *noch einmal*, this time as a student 1959-60.

University of California, Los Angeles (UCLA) for graduate work in psychology, a discipline less than a century old and still trying to define its subject matter. However, before I was enrolled in the graduate program, I received a Fulbright fellowship for a year of study in Germany. I was sent to the Medical Academy in Düsseldorf (now part of the Heinrich Heine Universität Düsseldorf, founded in 1965, a few years after I left). The Medical Academy was not my choice, but like the Army, the Fulbright Commission sends you wherever there is an opening that seems remotely appropriate to your interests.

At the Medical Academy, I was assigned to work with a psychologist in the psychiatric institute. My major task was to score some psychological tests to see if they might help identify patients who had suffered possible brain injuries on the job, even though they showed no signs of neurological damage. Workmen's compensation depended on an accurate diagnosis. Not surprisingly, the results of this exploratory study were inconclusive. In the fifty years since that time, techniques for neurological and neuropsychological assessment have advanced considerably, but are still far from definitive. So, a basic conundrum remains: Unmoored from identifiable (and potentially treatable) organic processes, suffering becomes like many other psychological phenomena, sensitive to rewards and punishment. Since monetary compensation can serve as a "reward" for suffering, a vicious cycle may be set in motion: the greater the suffering the greater the compensation (reward), and vice versa. How can suffering be alleviated when compensation is linked to its intensity and continuance? Like any good conundrum, this one has no simple solution. It is, however, an issue that calls for attention from a social and not just a psychological perspective; that is, we must take into account the institutions (especially the legal and health-care systems) that may benefit from the woes of individuals, and also cultural influences, such as romanticism, that elevate suffering to a status symbol (Averill, 1989).

For my second Fulbright semester I transferred to the University of Bonn, where the International Congress of Psychology was scheduled to be held. At Bonn, I was most influenced by a laboratory course in ethology (a branch of zoology), observing the behavior of

bees. I did not find the bees particularly interesting, but I did learn a valuable lesson: To understand behavior it is important to know how the behavior fits into a person's everyday life. Of course, humans cannot be observed constantly as they go about their daily business, as can bees. However, surveys, self-reports, and like methods can provide useful information, especially when supplemented with experimental studies on component processes.

Upon my return to UCLA, I planned to enter a program in Clinical Psychology. I found the problems fascinating, and still do. However, I soon realized that I have neither the patience nor the temperament to be a good clinician. Fortunately, after my first year, I was offered a four-year fellowship with only one stipulation, that I pursue a degree in biopsychology, or what was then commonly called physiological psychology. It was an opportunity I couldn't resist although it involved some major retooling, for I had little background in physiology and related disciplines. The fellowship also provided the time to continue my interests in philosophy, with an emphasis on "analytic" or "ordinary-language" philosophy.

You got your PhD at the University of California Los Angeles. Who were your mentors and role models back then? What were the topic and the main results of your dissertation?

My primary sponsor at UCLA was Marion Wenger, a major figure in psychophysiology, which uses physiological measures to evaluate psychological processes (as in lie detection). Wenger was a meticulous researcher. At that time, however, he was approaching retirement and was heavily involved in University administration. I seldom saw him personally and worked only occasionally in his laboratory. Still, I believe I owe much to his behind-the-scenes support. And I did my dissertation on a problem closely related to his concerns.

Wenger's psychological research focused on the emotions, which he conceived of as patterns of activity in organs innervated by the autonomic nervous system (Wenger, 1950). In this, he followed the lead of the Danish physician, Carl Lange (1885/1922). In contrast to William James, with whom Lange's name is often associated (cf.

the James-Lange theory of emotion), Lange conceived of emotions as responses of the cardiovascular system. "Is it possible", he asked rhetorically, "that vasomotor disturbances, varied dilation of the blood vessels, and consequent excess of blood, in the separate organs, are the real, primary effects of the affections, whereas the other phenomena, — motor abnormalities, sensation paralysis, subjective sensations, disturbances of secretion and intelligence — are only secondary disturbances, which have their cause in anomalies of vascular innervation?". Lange answered his question affirmatively, with numerous hypothetical examples. James (1884) framed his theory somewhat differently: "My thesis," he stated, "is that *the bodily changes follow directly the perception of the exciting fact, and that our feeling of the same changes as they occur IS the emotion*" (emphasis in the original).

These two formulations might seem very similar. James evidently thought so, for when revising the original 1884 publication of his theory for inclusion as a chapter in his *Principles of Psychology* (1890), he quoted long passages from Lange as support. But there is an important difference between the two. How the basic concepts of a theory are defined is fundamental to the further development of the theory. For Lange, an emotion is a vasomotor response, all else is secondary, including "subjective sensations". For James, an emotion is a *feeling*, a perceptual experience, albeit of bodily changes. "Let not this view be called materialistic", James admonished his readers. Although Lange was not as explicit on the topic, I doubt that he would have objected to his view being called materialistic.

In a sense, James and Lange stand at a crossroad on how emotions might best be defined. At the crossroad, the view ahead may seem similar. But once one road is chosen over the other, the landscape may quickly change. James's conception of emotions as feelings leads in a phenomenological direction; Lange's, in a physiological direction. Wenger chose to follow Lange. The major difference was that Wenger included in his definition of emotion all responses mediated by the autonomic nervous system (both sympathetic and parasympathetic branches), not just vasomotor responses.

I emphasize this point because I have proposed a still different definition of emotions,

namely, as transitional social roles. Initially, at least, this conception does not diverge greatly from the paths chosen by James and Lange. Down the road, however, the landscape quickly changes. But I am getting ahead of myself. Let me turn to the second part of your question, concerning the topic and results of my dissertation.

One implication of the James-Lange-Wenger tradition is that each emotion involves a unique pattern of visceral responses. Testing this implication had been the topic of considerable debate and research during the first half of the 20th century, but without resolution, due, in part, to the narrow range of emotions sampled (primarily fear and anger, both of which may involve strenuous physical activity) and the limited number of physiological responses typically assessed. For my dissertation, therefore, I compared two very different emotional experiences, sadness and mirth, on a variety of autonomic response variables. There were differences which I will not discuss here, for they are not directly relevant to the topic of this interview (Averill, 1969). One finding, however, did start me on a line of investigation that ultimately led to a social-constructionist view of emotion.

During the sadness condition (viewing a film of the funeral procession of the recently assassinated President Kennedy), participants showed physiological arousal, characteristic of the “fight or flight” response, but without any discernable impulse toward physical activity, overt or covert. This does not mean that sadness is generally marked by physiological arousal. Much undoubtedly depends on the eliciting condition and whether the sadness is an immediate, short-term reaction to events, or a more enduring response to a chronic condition. In the context of my dissertation, “sadness” could just as well have been described as a “mild grief response”. However described, a consideration of sadness leads almost ineluctably to the larger question of grief.

Many of the behaviors of individuals during grief seem paradoxical from a biological point of view. For example, a bereaved person (following, say, the death of a loved one) may not only show a high degree of arousal, as in acute sadness, but also may withdraw from social contact, refuse to eat, engage in self-mutilation, be susceptible to

disease, and much more. From a biological perspective, it would seem more fitting for the bereaved to return to normal social activities as quickly as possible. Why does the opposite seem to be so common?

Avoidance of pain is a stronger motivator to action than is the prospect of a reward. Grief is a painful reaction to loss (e.g., separation from a loved one). Most losses are not irrevocable; mitigating the pain of grief may thus be an important incentive to repair a loss. When repairing a loss is not possible, as when a loved one has died, grief reactions may nevertheless run their course, even to the detriment of the individual (Averill, 1968). In other words, grief may in fact serve an important evolutionary function, namely, maintaining social bonds by making separation a painful experience. Grief becomes dysfunctional primarily when occasioned by an inappropriate occurrence (such as the loss of a pet canary) or excessive in degree or duration. Of course, what is considered inappropriate or excessive depends, in part, on individual circumstances and social custom.

The above ideas, I soon discovered, were not original on my part. John Bowlby (1961) and David Hamburg (1963) had previously reached similar conclusions, from psychoanalytic and biological perspectives, respectively. Which brings me to a tangential but important lesson I learned from my studies of sadness and grief: If ever you believe you have an original idea, a few hours in any good library will typically disabuse you of any such conceit. Of course, your library search will not be random; you must have some vague idea where to look, probably from sources you no longer recall explicitly. Einstein captured the situation succinctly when he supposedly observed: “*Originality is forgetting your sources*”. (Variations on this theme have been attributed to a many different authors; its actual origins seem to have been forgotten.)

I learned another important lesson from the studies of sadness and grief. I originally made a theoretical distinction between grief as a biological reaction to separation or loss, and mourning as socially prescribed responses to bereavement (Averill, 1968). For example, some symptoms of grief, such a loss of appetite for food or sex, sleep disturbances, and even self-mutilation, occur in infrahuman animals,

particularly primates, where the maintenance of social bonds between individuals (e.g., epitomized by the reciprocal attachment between an infant and its care giver) is an important biological adaptation. Other responses, including elaborate funeral practices, are social in origin and vary widely from one culture to another. Upon further consideration, however, I concluded that such a distinction, while it may make sense in the abstract, does not do justice to the actual experiences of bereaved individuals. From a subjective (phenomenological) point of view, socially based mourning practices may be experienced as compelling (“instinctive”) as any biologically based response.

In addition to whatever biological functions grief might serve, it also has important social and psychological functions (Averill, 1979). For example, the death of an important cultural or political leader may be used to reinforce group loyalties even among persons who may have disagreed with, or seldom even thought about, the deceased. As the great French sociologist, Emile Durkheim (1915) observed: following a cruel loss, “One weeps not simply because he is sad, but because he is forced to weep. It is a ritual attitude which he is forced to adopt out of respect for custom”. This is an over simplification, perhaps, but for an individual experiencing a meaningful loss, does it really make a difference whether he or she weeps as a biological necessity or out of respect for custom? And how would you tell the difference?

You began your academic career as a lecturer at the University of California Santa Barbara in 1965, then became Assistant Research Psychologist at UC Berkeley in 1966, and finally moved to University of Massachusetts Amherst in 1971, where you have taught until your retirement in 2006. What made you elect the University of Massachusetts as your life-long academic home?

I went to UC Santa Barbara as a one-year replacement for a faculty member who was on sabbatical. During that year I also completed my dissertation. I then accepted a research position with Richard Lazarus at the University of California Berkeley. Lazarus had been asked by the U.S. Rehabilitation Services Administration to establish a five-year program to train graduate

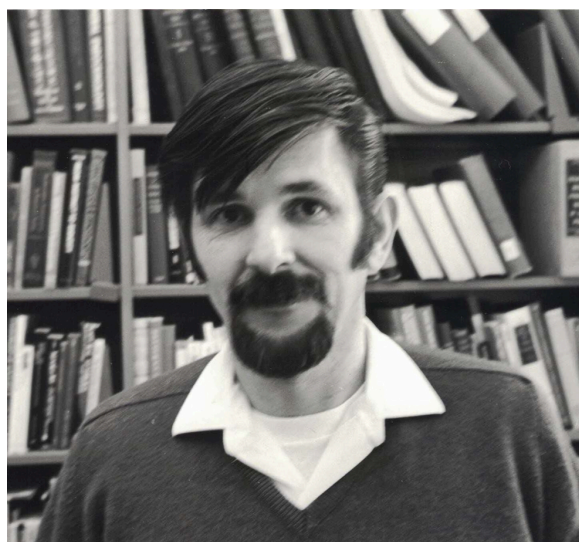
students to do research related to rehabilitation, for example, recovery from cancer, stroke, and spinal injury. He agreed, provided he could hire an assistant to help with the day-to-day operation of the laboratory, which involved psychophysiological recordings. I became that person. For me, it was like an extended post-doc, but with better pay and benefits.

Earlier, you asked me about role models I may have had, as well as mentors. I had many fine mentors at UCLA, whom I respected but did not identify with on a personal level. On the other hand, Richard Lazarus was less a mentor and more a role-model, at least as I conceive of the difference. He was highly disciplined in his work habits and a prolific writer, characteristics that can be modeled but not taught. As much as I would have liked to emulate him in these respects, that was not possible. I am constitutionally disorganized, and writing has always been a bit of a chore for me. Lazarus served as a role model in another respect, and, I hope, more successfully. In spite of keeping a very busy schedule himself, he always found time to give to others. As head of the Laboratory on Stress and Coping in which I worked for five years, his influence was often subtle but pervasive. If I were to describe in a few words the atmosphere of the lab, they would be “nonauthoritarian”, “collaborative” and “free.” In addition to common goals, Lazarus encouraged those who worked with him to pursue their own interests and ideas. That is an atmosphere I have always tried to create with my own students.

On a more intellectual level, Lazarus also had an important but indirect influence on my approach to the emotions. He is perhaps best known for his “appraisal theory”, which focuses on the way a person evaluates a situation as central to stress and coping. While I agree with that focus, it is not central to my own theorizing. Lazarus was trained as a clinician and his primary emphasis was on psychodynamics, that is, on the ways that individuals cope with threatening situations. My own interests were more on group than individual processes, whether the group was a species (evolutionary psychology) or a society (history and culture). That difference in emphasis, although minor in the short term, leads in different directions in the long run. Anyway, Lazarus’s emphasis on appraisal paved the way

for a more cognitively oriented approach to emotion (Lazarus, Averill, & Opton, 1970), which, in turn, opened the door to considerations (including social constructionism) not traditionally associated with the study of emotion.

When the grant from the Rehabilitation Services Administration ended after five years, I accepted a position as Associate Professor at the University of Massachusetts (UMass) Amherst, where I remained until retirement. The move from Berkeley to Amherst required some adjustment, especially with regard to the cold winters, but I could not have asked for a warmer or more supportive environment than UMass in which to pursue my academic interests. Moreover, Amherst is a lovely town in which to raise a family (I had two preschool daughters at the time). Amherst is located in the valley of the Connecticut River, about 100 miles west of Boston and 200 miles north of New York City. Education is the major “industry” in the area. Amherst College, Hampshire College, Smith College, and Mt. Holyoke College (four private, distinguished undergraduate institutions), and UMass Amherst, the flagship campus of the University of Massachusetts system, are all in commuting distance of one another. Each maintains a schedule of invited speakers, as well as first-class musical and theatrical events. The town of Amherst is thus unique in having many of the cultural advantages of a city, while being in a semi-rural setting.



Jim Averill, UMass 1974.

You became a member of ISRE at its founding in 1984/5. What are your earliest memories of it? Do you have a favorite ISRE conference? Do you like the way the society has been developing in the past few years, and do you have any advice concerning ISRE's future?

I do not have a favorite ISRE conference, although I did enjoy earlier conferences more than later ones. When ISRE was first formed, the number of people doing research on emotion was relatively small and qualifications for membership were strict; for example, all members had to have a record of research and publication in the area of emotion, and new members were subject to review by a membership committee. As a consequence, most members of ISRE were personally acquainted; conferences were small, often located in a college dormitory or similar facility; parallel sessions were not scheduled, so everyone could attend all presentations; and ample time was set aside for informal gatherings. Such an arrangement had many advantages, but also serious limitations. Most important, it tended to exclude young investigators just starting their careers and discouraged even established investigators working on the fringes of what is now commonly called “affective science”. As an organization ISRE thus ran the risk of becoming sclerotic and increasingly irrelevant in a rapidly expanding field.

But expansion has also presented difficulties. There are only so many conferences an individual can attend, because of time and expense. Physiologists and neuroscientists, who have always played an outsized role in emotion research and theory, have organizations better suited to their specialties. The same is true of sociologists and anthropologists. As a consequence, ISRE has become less interdisciplinary than originally envisioned. There is no need to call attention to this problem, for it is well recognized. Unfortunately, there is no simple solution. It is up to the present membership of ISRE to decide what kind of organization they want and how best to achieve it.

However, I do have several ideas that might be worth mentioning. To help preserve both the international and interdisciplinary aspects of ISRE, perhaps adjunct and honorary categories of

membership could be established and committees appointed with the goal of recruiting worthy candidates. Also, as important as ISRE conferences are in facilitating interpersonal contacts, I believe its publications are even more important. The *Emotion Review*, under the leadership of Jim Russell, Lisa Barrett and Christine Harris has become an excellent journal. Rather than simply waiting for submissions, Jim, Lisa, Christine and their “Special Section” editors have solicited target articles and commentaries from experts both within and outside of ISRE, with an emphasis on potential future as well as past developments..

You started your career as a physiologist of emotion in the 1960s, and in the mid-1970s you became one of the founding fathers of the modern social constructionism movement. Can you pinpoint to the events that led your perspective on emotion to shift from the physical to the social? In other words, why did you come to think that the social dimension of emotions was important and neglected?

My first publication (Averill, 1980a) presenting in detail a social-constructionist (then labeled “constructivist”) view of emotion may have seemed like a radical departure from generally accepted theories of emotion. But I did not view it that way. It had its origin in my attempt to achieve a unified view of grief and mourning, discussed in response to an earlier question. Also influential was my work with Dick Lazarus on the role of appraisal in stress and coping, mentioned earlier. And, of course, there was ample precedence in the work of others. Particularly noteworthy was Ted Sarbin’s analysis of hypnotic trance as a form of role-playing; this had a direct influence on my own interpretation of emotional states as transitional social roles. And no discussion of social constructionism would be complete without reference to Ken Gergen’s extensive and erudite publications on the topic.

Still, the path to a social-constructionist perspective was not without obstacles. Along the way, a major question had to be addressed: Why had the social dimension of emotions been so long neglected or, when recognized, relegated to a secondary role, as a mere patina on the underlying *real* emotion? Surely, so many people

for so long a time could not be wrong. But they were, I came to believe, and it is important to understand why. A key to that understanding is to be found in a phenomenon I call *psychophysiological symbolism* (Averill, 1974), that is, the association of psychological processes with physiological structures and activities on the basis of shared symbolic meanings. Psychophysiological symbolism plays a particularly important role in the localization of function when actual physiological evidence is weak. For example, the ancient Greeks had little knowledge of brain function. Therefore, it seemed reasonable to Plato to localize rationality in the head. Rationality, he believed, involved the circular movement of thought. Rationality being the highest manner of thought and circularity being the most perfect kind of motion, it seemed reasonable to Plato that rationality should be localized in the head, the part of the body most spherical in shape. And what about the emotions which, in Plato’s view, often interfere with rationality? He located them in the torso, separated from the head by the neck, a narrow passage, so that they might interfere as little as possible with rational thought. And the less desirable an emotion, the lower in the body it belonged, the basest emotions being localized below the midriff.

Plato had other reasons for his localization of function (e.g., the head is closer to the heavens), but they were equally symbolic. Plato’s theorizing in this regard seems quaint by today’s standards, but only because we know more about the actual functions of different organs in the body. Still, from the ancient Greeks to the present day, psychophysiological symbolism has played an important role in how we conceive of the relation between emotions and bodily functions. For example, emotions are generally considered involuntary (e.g., we are “gripped”, “seized”, and “overcome” by emotion); hence, it seems reasonable that emotions be associated in some way with the “involuntary” (autonomic) nervous system. To take another example, emotions are often described as “brutish”, “bestial”, and the like; hence, it seems only reasonable that emotions are remnants of our biological (animal) heritage, mediated by phylogenetically more primitive parts of the brain, e.g., the limbic system and paleocortex.

In short, our body is as much symbol as substance; our emotions, too, are imbued with symbolism. When a coincidence of symbolism exists between body parts and emotions, it seems reasonable to assume the two are related. And when the relation gets embedded in our ordinary language, the localization of function becomes self-evident and especially difficult to dislodge.

At the heart of your theory of emotions is the notion of the “myth of the passions”. Can you explain what the myth is, where you think the myth comes from, and why it is important to dispel it? What is the connection between the myth of the passions and the way we talk about emotions (e.g., the metaphors we use)?

I borrow the phrase, “myth of the passions” from Robert Solomon (1976), who contrasted it with a corresponding “myth of reason”. Both myths have their origins in classical Greek thought. For the Greeks, reason or rationality was the defining feature of human nature. By contrast, passions (a generic term that included what we now call emotions) were believed to interfere with reason, often to the detriment of the individual or society.

In response to the previous question, when discussing psychophysiological symbolism, I suggested one way in which the concept of emotions as passions has influenced theories of emotion. At that time, I was not familiar with Solomon’s work. His phrase, “myth of the passions”, is particularly apropos in highlighting how the concept of emotions as passions has had a pervasive influence on theories of emotion, beyond psychophysiological symbolism. And it suggests why that influence has been so resistant to change. Myths are not ordinary beliefs; they are beliefs we live by, even when we recognize that they have little basis in fact. To take a trivial, everyday example, I do not walk under ladders, not because I think it brings bad luck, but because it is a myth I act on without thinking. Not all myths of the passions are as trivial or inconsequential as this example might suggest.

The distinction between passions (things that happen to us) and actions (things we do) was given its most influential formulation by philosophers like Plato and Aristotle; however, even before their time the distinction was deeply embedded in the Greek language and thought.

Specifically, *pathē* was the Greek term for what we now call “emotion”. *Pathē* could refer to any object, animate or inanimate, that was undergoing (“suffering”) change through the action of an external agent. A piece of wood, for example, could suffer the blow of an ax. Emotions were one category of *pathē* that a person might suffer; diseases were another. Hence, from *pathē* we get such medical terms as “pathology,” “pathogen,” “idiopathy,” as well as such emotional terms as “pathetic,” “empathy,” and “antipathy.”

The Greek *pathē* was incorporated into Latin as a form of the verb, *pati*, which also means “to suffer”. In the transition, *pathē* lost some, but not all, of its pathological connotations. The past participle of *pati* is *passus*. Hence, the emotions, or at least the more turbulent emotions, came to be known as “passions.”

Dixon (2003) has traced the use of the term “emotion” as it came to be preferred over “passion,” first among the Scottish moral philosophers (e.g., Thomas Brown) and later by such theorists as Bain, Spencer, Darwin, and James. The switch in terminology from “passion” to “emotion” did not, however, eliminate the connotation of passivity (“suffering”) that marked the basic concept for the preceding two millennia, nor did it do much to debunk the Myth of the Passions. On the contrary, according to Dixon, the contemporary concept of emotion has lost many of the subtleties associated with earlier analyses of the passions; hence, if anything, the myth has only grown stronger because it is less evident.

A conceptualization of emotions as passions continues to have a profound influence on the way we talk about emotions in everyday life. It may help explain, for example, why emotional concepts with a negative connotation outnumber positive concepts by a ratio of about two to one (Averill, 1980b). After all, who wants to suffer gladly?

As implied by your question, the “myth of the passions” could also be labeled, “passions as metaphor”. That rephrasing highlights an important point. Metaphors have played an important role in theories of emotion (Averill, 1990), and in science generally. For the most part, that role has been positive. For example, during the late Middle Ages the conception of the universe as a huge clock-like mechanism, with

God as the master clockmaker, helped advance the Scientific Revolution. However, by making some lines of thought seem natural and hence unquestionable, metaphors can also have a stultifying influence. Because its implications are so broad, I believe the myth of the passions has had more of a negative than a positive influence on theories of emotion.

Constructionism is all the rage these days. What are the main tenets of your own version of social constructionism about emotions?

“Social constructionism” is often contrasted with “realism” as a philosophical doctrine. Realism comes in many varieties; what the varieties have in common is the assumption that reality, however conceived, exists independent of human thoughts and desires. Put most simply, reality is to be discovered, not made. Social constructionism, by contrast, asserts that reality is always relative to the human condition at a specific time and place.

On a practical level, social constructionism emphasizes the role of language (discourse) as a tool for constructing the realities we know, or think we know. Of course, language is not the only tool in the social-constructionist tool-box. Art, music, ritual, even technological innovations, are also important. But language is fundamental.

An emphasis on language is not specific to social constructionism. As the existentialist Martin Heidegger observed: “Language is the house of Being. In its house man dwells”. Heidegger could not be considered a social constructionist. On the contrary, a main goal of his philosophy was to escape from the house of language and discover the true nature of Being (*Dasein*). For the social constructionist, that is a vain attempt. However that may be, my concern is primarily with the emotions, and the contrast is not with reality (or Being) in general, but with biological and psychological determinism; that is, with the notion that emotions exert an influence on behavior independent of, or contrary to, a person’s reason or will. I leave the broader implications of social constructionism for others to debate (Parker, 1998).

Specifically, with regard to emotions, a social-constructionist view rests on three assumptions: first, emotions are complex

syndromes (systems of behavior) comprising diverse, semi-autonomous components; second, no one component or class of components — physiological, behavioral, or cognitive — is essential to the whole; and third, social beliefs and rules are the primary principles by which the various components are organized into wholes.

The first assumption (a componential approach) is becoming increasingly common among emotion theorists of diverse persuasions. It is based on the recognition that emotional terms such as “anger”, “love”, and “fear” do not refer to specific responses, but to *syndromes*. Briefly stated, an emotional syndrome is a set of interrelated responses, including an appraisal of the situation (what the emotion is about), the expected outcome (objective), and the way the whole — the manifest thoughts and responses — are experienced (interpreted), namely, as a passion (something that happens to us) rather than an action (something we do). When emphasis is placed on social influences or “rules” as the primary source of coherence (mutual interaction among the elements of a syndrome), we may speak of the social construction of emotion. This is in contrast, say, to disease syndromes where the interactions among symptoms is due primarily to physiological principles, such as homeostasis.

The second assumption (non-essentialism) also has ample logical and empirical support. Without going into detail, emotional concepts cannot be defined “classically” (in terms of necessary and sufficient conditions). Rather, they refer to conditions that are related by what Wittgenstein called “family resemblances”; that is, overlapping features none of which need be shared by all members of the group.

This brings me to the third and most controversial assumption underlying a social-constructionist approach, namely, social beliefs and rules are the primary principles by which emotional syndromes are organized and interpreted. I presume we will return to this issue in more detail in response to future questions. I will therefore conclude these introductory remarks by drawing a few contrasts between a social-constructionist approach, as I conceive of it, and more traditional theories of emotion.

All theorists recognize that emotions are regulated by “display” rules, for example, in most

cultures it is considered inappropriate to laugh at a funeral. Less commonly appreciated is that many rules have a *constitutive* as well as regulatory function. To illustrate what I mean by a constitutive function, let us consider a non-emotional example: the rules of English grammar help regulate how to speak properly, for example., in well-formed sentences that are easily understood. More fundamentally, English grammar helps make (constitute) the language what it is: English as opposed, say, to German or Chinese. Simply put, without the rules of English grammar there would be no English language to regulate. Similarly, without the rules of anger, say, there would be no anger to regulate, only inarticulate rage reactions, or perhaps some other socially constructed emotional syndrome (such as envy or jealousy).

You distinguish between emotions as episodic dispositions, cognitive schemas and transitory social roles. Can you say more about these distinctions?

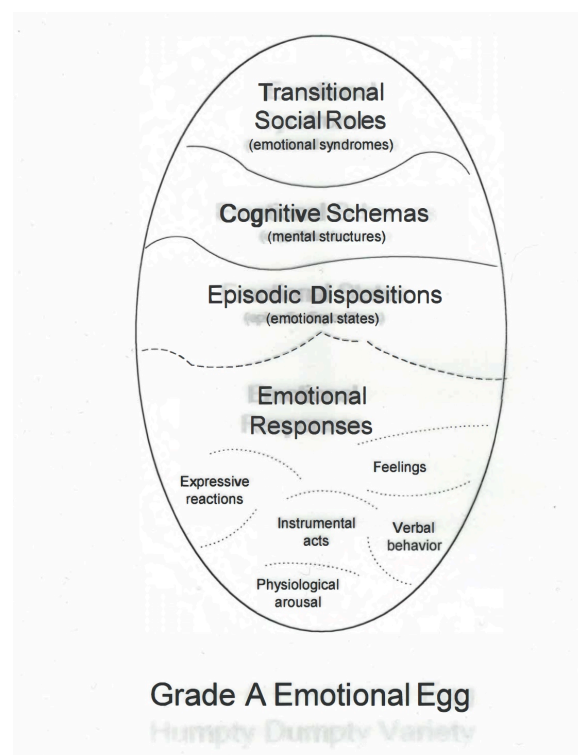
It is a commonplace expression that you have to break an egg to make an omelet. Social constructionism is not about making omelets; it is, rather, about the living, clucking chicken that develops from an egg under normal conditions. Episodic dispositions, cognitive schemas, and transitory social roles are three overlapping features or principles that help us understand how emotional responses are organized into coherent syndromes. Any discussion of one presumes the other two. Nevertheless, they point in somewhat different directions for further analysis, as I will try to illustrate briefly.

The phrase *episodic disposition* refers to the emotional state a person happens to be “in” at the moment. Stated more formally, an emotional state is a short-term, reversible (episodic) tendency (disposition) to respond in a manner characteristic of an emotional syndrome. This is a simple but often misunderstood formulation. The misunderstanding stems mainly from the manner in which emotional episodes are often identified, namely, by letting part of a syndrome stand for the whole. For example, “sweaty palms” is commonly used as a stand-in for fear; a frown, as a stand-in for anger; and so on for other emotional syndromes. But as already explained, fear can be expressed in a great variety of ways,

as can anger, and most other emotional states.

At the risk of being repetitious, let me reiterate what I have already said, namely, that emotions (conceived in the abstract) are *syndromes*, not specific responses, and no single response or type of response counts as a sufficient or necessary manifestation of the whole. This means that when a person is in an emotional state (episode), he or she may respond in any of a variety of ways consistent with the emotional syndrome. To account for the diversity of ways in which an emotion may be expressed, an emotional state can best be conceived of as an episodic disposition: It is “episodic” because the state is temporally delimited, typically with a specifiable onset and offset; it is a “disposition” because being in an emotional state is not per se a response, but only a readiness to respond, or not respond, depending on the person and situation.

To further illustrate this point, consider the distinction commonly made between emotional states and traits. Both are dispositions, but they may differ both in duration (short-term vs. long-term) and breadth (a narrow vs broad range of potential responses). For simplicity, I will focus here primarily on the temporal dimension.



First, consider trait-anxiety, one of the most frequently assessed dimensions of personality (often labeled “neuroticism”). It is a relatively enduring disposition (it can last a lifetime) to perceive and respond to situations as threatening. Contrast this with state-anxiety, which is a short-term, reversible (episodic) disposition to respond at the moment to a threatening situation. Emotional states, because they are episodic dispositions, are often limited to only a few modes of expression (sometimes called “occurrent responses”). As already mentioned, this restriction of responses has often led researchers to identify an emotional state with one of its prototypic manifestations; then, generalizing from emotional states, emotions as a category may be identified with only one or a few responses. Carl Lange’s identification of emotions with vasomotor responses is a good example of this type of generalization.

Let me turn now to *cognitive schemas* as they relate to emotions. By “schema”, I refer to mental structures (e.g., beliefs, concepts, goals, and plans) on which emotional appraisals are based, responses organized, and stimuli and responses connected. I also interpret “cognitive” broadly to include both emotional (“hot”) and intellectual (“cold”) schemas.

Some schemas are shaped by our evolutionary past; others are a legacy of our culture; still others are a product of individual experience; and most are a combination of all three sources, in varying degrees. Whatever their source, without cognitive schemas, the world in which we live in would be, to paraphrase William James, a blooming, buzzing confusion.

The study of cognitive schemas belongs to many disciplines. For example, identifying the relevant dimensions along which emotional schemas may differ from intellectual schemas, and how emotional schemas relate to specific emotional syndromes, are the subject matter of “appraisal theory”. The way responses are organized and interpreted has also been a traditional concern for emotion researchers, both physiologists and behaviorists. And between the appraised object and an emotional response, many intervening steps may occur (cf. Freudian defense mechanisms, such as repression, sublimation, projection, reaction formation, and the like). More recently, computer scientists have

gotten into the act, exploring how an affective component might be added to their algorithms, thus making problem-solving more effective, even if less rational.

In short, there is no common terminology to describe cognitive schemas, no less an integrated theoretical approach. I will therefore leave the topic at the level of ordinary language; that is, schemas are the “concepts”, “beliefs”, “goals”, and “plans” by which we make sense of experience and guide behavior. The only thing I would add for emphasis is that schemas are dynamic, not static structures; they may undergo continual and dramatic change during the course of an emotional episode. To say more at this point would add more jargon than clarity to the discussion.

I want to turn now to the third part of your question: What is added to dispositional and schematic analyses by referring to emotions as *transitory social roles*? Or, as I now prefer to phrase it, *transitional* social roles, for the role itself is not transitory, only the time people engage in the role as they transition from one nonemotional state to another.

Both dispositional and schematic analyses refer to aspects of the individual. Indeed, schemas are one way of characterizing the structural variables that make emotional dispositions possible. With few exceptions, however, emotions are social phenomena. The concept of emotions as transitional social roles helps to bridge the gap between the individual and the social, and it raises issues that can only be addressed on a social level of analysis. Chief among the issues associated with social roles are *legitimization*, *privileges*, *restrictions*, and *obligations*.

But before getting to that, let me say a few words about social roles in general. We are all familiar with the concept of a role from its use in theater. An actor who plays Hamlet in Shakespeare’s play by that name is, redundantly speaking, “playing a role”, that is, following a script and other stage directions. The notion of a social role builds on this idea, except the stage is the world in which we live and the script is written by society, albeit with ample opportunities for individual improvisation. Shakespeare also recognized this point when he wrote in another play: “All the world’s a stage,

and all the men and women merely players; they have their exits and their entrances, and one man in his time plays many parts". He then went on to describe an emotion that tends to dominate each of seven ages of a person's life, for example, during the third (young adulthood) stage, we find "the lover sighing like a furnace" (*As You Like It*, Act II, Scene VII).

Currently, to say that someone is playing a role has a mildly pejorative connotation, as though the person is not being authentic. That connotation does not necessarily carry over to social roles. Indeed, it does not always apply even to theatrical roles. Accomplished "method" actors, for example, may become so engrossed in the roles they are playing that the script temporarily constitutes their reality.

Broadly defined, role-playing is a form of rule-following behavior. To illustrate, consider the role of a judge in a court of law. To engage in that role, a person must know the law and follow the rules of behavior expected of a judge. But that is not enough. A delusional schizophrenic might play the role of a judge, following all of the rules relevant to the role and sincerely believing that he is a judge. Yet, he would not be a judge. Engagement is a social role also requires *legitimation*, a concept that has meaning only on the social level of analysis. To gain legitimacy, a person engaged in a social role generally must meet certain entry requirements, such as having passed the bar exam in the case of a lawyer or judge.

Similar considerations apply to emotional roles, although the entry requirements are seldom as obvious. A good example is the legal requirement for attributing a homicide to anger, thus mitigating the charge from murder to manslaughter (a "crime of passion"). Not only must the defendant have conformed his or her behavior to relevant rules, but the jury must also confer legitimacy on the performance by having it pass the so-called "reasonable man test" (about which I will have more to say shortly). A more familiar example might be the case of an adolescent who becomes infatuated with a youngster of the opposite sex. The emotion is not likely to be legitimized as true love, not because the couple lacks sincerity, but because the individuals do not occupy an adult social status, and hence cannot fulfill all the obligations of the

role. Of course, customs change. In Shakespeare's time, the relationship between Romeo and Juliet might have been considered a paradigm of love; today, the couple might be treated as juvenile delinquents, had they not killed themselves first.

In addition to questions of legitimation, social roles are associated with privileges, restrictions, and obligations. Emotional roles are no exception. Brief references to anger, love, and grief will suffice to illustrate the point. Extrapolation to other emotions is relatively straight-forward.

Privileges. Emotional roles allow a person to engage in behavior that would be discouraged under ordinary circumstances. As just described, for example, an angry person can literally get away with murder (i.e., have a charge of homicide mitigated from murder to manslaughter). When in love, a couple may engage in sexual behavior that might otherwise be discouraged. And while grieving, a person may be exempted from obligations related to work and entertainment.

Restrictions. Privileges have their limits; a person can only do so much when emotional and "get away with it." In spite of presumably being beyond personal control, emotional responses should be appropriate to the situation: They should not be too mild or too intense, too short or too prolonged, or too idiosyncratic. For example, a plea of anger will not be accepted in a court of law if the crime is committed in too cruel or unusual a manner. Similarly, lovers are expected to be discrete and honorable in their affairs, and their liaison should last for more than a few hours. And if a bereaved spouse begins dating very soon after the death of a partner, the genuineness of his or her grief may be called into question.

Obligations. Whereas there are some things a person cannot do while in an emotional state (restrictions), there are other things that should be done. An angry person, for example, is expected to take action to correct the appraised wrong, or else the sincerity of his anger, or even his character, may be questioned. Love, too, has its obligations, for example, a commitment to the well-being of the other. And, with regard to grief, the bereaved person who fails to comply with socially prescribed mourning practices may be subject to severe sanction.

In sum, conceiving emotional syndromes as

transitional social roles adds a new dimension to their analysis. Not only do emotions involve episodic dispositions to respond, mediated by cognitive schemas, they also have irreducible social aspects having to do with legitimation, privileges, restrictions, and obligations.

Some emotions do not seem to fit especially well your view that emotions are transitional social roles interpreted as passions, as there doesn't seem to be anything social about them. Just to pick an example, it is hard to understand how the sort of fear one experiences when suddenly losing support may amount to a social role. Should we then conclude that social constructionism is a theory that sheds light on some rather than all emotions?

That is a reasonable conclusion. Emotions form a heterogeneous category and no one formulation is equally applicable to all affective states. Social constructionism is least applicable to (a) reflex-like reactions, such as fright experienced on the sudden loss of support, and lashing out at a source of pain; and (b) broad affective dispositions, such as states of undirected excitation, anxiety, and depression. The relevance of a constructionist view increases with the complexity of the emotion and the involvement of cognitive mediating mechanisms (schemas). Thus, as important as the above exclusions may be, a social-constructionist view is applicable to the majority of discrete emotions recognized in ordinary language.

Having said that, I would like to add a qualification: from a social-constructionist perspective, it is a mistake to take the simplest manifestation of an emotion as representative of the general category, for example, fear of falling as representative of fear in general. In one sense, fear of falling is a prototypic fear, especially among very young children and the elderly. But it is not a typical fear for most people. In contemporary American society, fear of terrorism (and its derivatives, Islamophobia and xenophobia) has a far greater influence on behavior than fear of falling, even though very few people have been injured by terrorists (especially relative to the tens of thousands injured in automobile accidents each year). The role of society in many fears is particularly

evident from a historical perspective. For example, if we go back a few centuries in most Western societies, anyone who did not manifest a fear of God, and act accordingly, was likely to be roundly condemned; in the extreme, he or she might even be executed as a wizard or witch.

In short, most fears are thoroughly saturated with social beliefs and rules. It might be argued that societies only use a primordial fear for their own purposes, by attaching it to socially relevant objects and socially constructed responses: that somewhere between the eliciting condition and sanctioned response, a *basic* fear exerts its influence. Unfortunately, no one has been able to find or describe to everyone's satisfaction the nature of that basic fear.

In making this last observation, I do not wish to gainsay the excellent research done by Jaak Panksepp, Joseph LeDoux, and others on neurophysiology of a few "primary" or "basic" emotions, fear chief among them. It is worth noting, however, that the results of such research are abstractions, what is left over when individual experience and (in the case of humans) social and cultural influences have been subtracted. The extent to which such findings can provide the basis for a general theory of emotion is thus open to question. (Although I do believe they can provide important insights into some emotional disorders.)

You have written about several specific emotions, so I would like to ask you about a few of them. Let us begin with stress. What is your understanding of stress and why do you consider it to be an important emotion to understand? What are the best ways to regulate stress?

The phrase "stress and emotion" is commonplace, implying that the two are conjoined. But in what way? On one interpretation, stress is another emotion among many, albeit one that is vaguely defined. An alternative interpretation is that stress is a generic category that includes emotions as members, much as a forest includes trees. In the following discussion I will focus primarily on this latter interpretation. The choice of examples (stress as a generic category or as a specific emotion) will be primarily a matter of suitability to illustrate a point.

The concept of stress has long been used in physics to describe changes that may occur in a material after it has been subjected to external forces, as when a metal becomes brittle after being subjected to repeated or continuous pressure. The stress concept was introduced into medical science by Austrian-Canadian physician, Hans Selye, who had observed that a standard set of physiological changes occur when an animal or human is subjected to any of a wide variety of potentially harmful events. Selye called this the General Adaptation Syndrome (GAS). The GAS begins with activation of the pituitary-adrenal hormonal system and the sympathetic nervous system. The purpose of this initial stage is to ward off the immediate threat and heal any injury that might have occurred. This is followed by a stage of resistance in which homeostatic balance is restored but resistance to further injury remains high. Eventually, if the stress continues, a stage of exhaustion may occur, during which the stress responses themselves may cause tissue damage, such as stomach ulcers and enlargement of the adrenal glands. Selye named these potentially dangerous consequences, “Diseases of Adaptation”.

In Selye’s original formulation, stress could be considered another tree in the emotional forest; the GAS is a well-defined syndrome, that is, a set of coordinated responses that develop in a predictable manner over time. However, Selye and others quickly saw the usefulness of “stress” to describe almost any response to a situation that exceeded a person’s, or animal’s, ability to cope. “Stress” thus came to be used as a generic term to include all kinds responses to taxing events (“stressors”), whether positive, as with the birth of a child, or negative, as with the death of a child. In other words, stress ceased to be another tree in the forest; it became the forest itself — and an inevitable part of life.

Central to this broader conception of stress is the notion of “appraisal”. As defined by Magda Arnold (1960), an appraisal is an intuitive recognition of an event as potentially beneficial or harmful, as when a sheep immediately recognizes a wolf as a potential danger. To this “primary” appraisal, Lazarus (1966) added the notion of “secondary” appraisal; in addition to recognizing threat, an appropriate coping response must also be determined. The terms

primary and secondary might suggest a temporal sequence, but that is not necessarily the case. As with perception in general, a focal stimulus is always interpreted within a context that includes potential responses, what Gibson (1979) has called “affordances”. For example, observing a bear behind a cage in a zoo is very different than meeting a bear while walking in the woods, in part, because the two situations afford different avenues for response.

Generally speaking, each commonly recognized emotion, with the possible exception of “free-floating” anxiety, excitement, and depression, is associated with a different kind of appraisal. To illustrate, ask yourself what distinguishes anger from envy, joy from gladness, shame from guilt? Responses, such as patterns of physiological arousal and expressive reactions, may provide part of the answer, but only part. The primary factor that distinguishes one emotion from another is the appraised object. Stress as a generic category is also associated with the way a situation is appraised, that is, as taxing or not. Stress being a generalized response, research has focused on such nonspecific stimulus variables as intensity, novelty, and uncertainty, as they might interact with individual differences in the desire for personal control (Averill, 1973).

A common misinterpretation is to assume that an appraisal is the *cause* of an emotion. For example, I am angry *because* I believe I have been unjustly criticized. But that is only partly true. As Dewey (1895) suggested in response to the question posed by William James’ (1884), *What is an Emotion?*, the way a person appraises a situation is part of, not antecedent to, the emotion. Some theorists go so far as to argue that emotions are in principle a kind of appraisal or evaluative judgment (e.g., Sartre, 1948).

As already noted, stress and emotions can be manifested in various ways, including appraisals, physiological arousal, expressive reactions, instrumental responses, and more. Once one component is activated, others may be recruited to complete the whole. That is part of what is meant when we speak of stress and emotions as *syndromes*.

Of course, not all components of stress and emotions are of equal importance, and of all the components, the way the situation is appraised is the most important. This fact lies at the heart of

many cognitive theories of emotions. On a more practical level, the centrality of appraisals means that the best way to short-circuit stress and emotions is to alter the appraisal, for example, by interpreting a presumed insult as a constructive criticism. But a benign reappraisal is not always possible or reasonable. Sometimes an insult is truly an insult; in which case, some other component of the syndrome might be altered. Breathing is especially important in this regard. Breathing exercises are a significant part of nearly every stress-reduction program. Not only does measured breathing help calm physiological arousal by assuring a healthy supply of oxygen to the brain and other organs, but a focus on each inhalation and exhalation as it occurs helps to concentrate attention on the here and now (“mindfulness” in the current argot of stress management).

You have suggested that emotions emerge when there are “norms which simultaneously encourage and discourage a particular kind of behavior”. For example, in your early work on anger you have argued that anger leads to the kind of behavior discouraged by norms against violence and encouraged by norms in favor of protecting one’s own rights from infringers. By being “overcome” by anger, individuals manage to protect their rights by inflicting violence and are justified in so doing so because anger allegedly overcame. Is this your theory of anger? How do you see the role of anger in the election of Donald Trump?

In making the above suggestion I was referring primarily to anger and, in particular, to anger as adjudicated in courts of law, where “adequacy of provocation” is one of the major criteria for legitimizing a response (homicide) as angry, that is, as a crime of passion. And what counts as an adequate provocation? An affront that would arouse a “reasonable man” to violence. Note that the homicide is still a crime, but one that is treated more leniently than a similar act committed with “malice aforethought” (murder). In other words, by committing homicide, the perpetrator upheld one set of norms (as validated by the “reasonable-man” test) but in doing so has violated another set of norms (against deliberately killing another person). The conflict is resolved by convicting the perpetrator

of a lesser offense, voluntary manslaughter, a crime of passion, rather than murder. In a sense, the “victim” (the one who provoked the angry response in the first place) is put on trial as well as the perpetrator (the killer), and is found guilty.

How far this line of reasoning can be extended to everyday experiences of anger, I will leave it to the reader to decide. I will only point out that, when a person becomes the target of another’s anger, the most common response is, “Why, what did I do wrong?” Not: “How do you know?” or “How does it feel?” If the anger proves justified, an apology or some other form of restitution by the target typically ends the episode (Averill, 1982).

Considerations that apply to anger may not apply to other emotions without qualifications, and in some cases not at all. Still, the use of emotions as an excuse for otherwise unwarranted behavior is more common than generally recognized.

You also ask: How do I see the role of anger in the election of Donald Trump? It is an interesting question. I believe it is too facile an explanation to attribute Trump’s election primarily to anger; more importantly, it short-circuits the need to ask more relevant questions. A reasonable assumption is that Trump voters, *like most other voters*, based their choice, not on anger, but on perceived self-interest. Postulating anger as a motive is akin to saying that a vote for Trump was an electoral “crime of passion”. Depending on one’s political orientation, that may be a satisfying description. However, it may only succeed in making Trump supporters angry, even if they were not so already.

In what sense is love socially constructed? Has our understanding of romantic love changed significantly over time? What would you say of a theory that proposes that love is just an evolutionary adaptation aimed at passing one’s genes to the next generation?

Making love is a common euphemism for having sex; yet, a couple can have sex without being in love, and can be in love without having sex. If that were not the case, the world’s oldest profession would have gone out of business long ago. Having said that, it is nevertheless the case that in Western societies love is one of prime justifications for having sex, which, until recent

years, was (ideally) delayed until after marriage.

But if not sex, what motivates love? Perhaps it will help to make a long story short if I draw a comparison between having sex and eating food. Without eating, a person will die within a matter of weeks; without sex, a species will become extinct within a generation. Not surprisingly, then, there are strong biological incentives to both eat and have sex. Yet, how the “hunger” for each gets satisfied varies greatly as a function of culture. Take eating. The food a person prefers is one of the markers of his or her ethnicity or social identity. For example, French are distinguished from Japanese, in part, by the way they prepare their food, even though from a biological (nutritional) standpoint Japanese and French cuisine may be equally healthy.

Now take romantic love (and hereafter when I refer to love, I mean the romantic variety as opposed, say, to parental or fraternal love). Are there cultural differences in love, as there are cultural differences in the way we eat? Most assuredly. A need for sex may be universal, but love transforms sexual desire into an emotional syndrome that meets social ends as well as, or even in place of, sexual satisfaction.

In many societies, procreation – begetting and rearing the next generation – is too serious a business to be left to the vagaries of sexual infatuation. For example, collectivist societies, where communal values are emphasized over individualistic values, may treat sexual infatuation as akin to a disease, of which a couple should be “cured” before they embark on the serious business of procreation. Individualistic societies have found a different solution to getting couples to marry and have children, even when cultural norms might encourage a more self-centered lifestyle. That solution is falling in love, a condition – like a crime of passion – that is ostensibly beyond individual control.

The Western ideal of romantic love is often traced to the courtly love of the middle ages, when a knight would pledge allegiance to a noble lady other than his wife. Sex between the knight and his lady was discouraged, at least in theory; hence, a wife (often wed for social or political purposes) was excluded as a potential object of courtly love.

Western societies have undergone many changes from feudal to modern times; so, too,

have conceptions of romantic love. In particular, the courtly ideal has changed from a nonsexual commitment to a partner outside of marriage, to a state that supposedly emerges between a couple after marriage, and, more recently, to a precondition for marriage. And love continues to evolve. No longer is love limited to couples of the opposite sex, and, contrary to the words of a popular song from the recent past, love and marriage no longer “go together like a horse and carriage”. Still, some of the memes that characterized courtly love (e.g., only one lover at a time, an idealization of the person loved, and willingness to sacrifice for her or him) continue to influence the experience and behavior of persons as they “fall” in love (Averill, 1985).

Are you religious? What role do you think emotions play in becoming religious?

According to the French mathematician and philosopher, Blaise Pascal (1623-1662), some people are “so made that they cannot believe”; to which Christopher Hitchens (2007) has added, “and there are more of us [nonbelievers] than the faithful would like to think” (p. 138). No, I am not religious, and the reason is, as Pascal insinuated, as much temperamental as intellectual. I was raised in an Irish Catholic family (on my mother’s side), and I attended parochial schools for six years. The theistic parts never “felt” right. No matter how strongly I tried to convince myself I should believe, I couldn’t. Nevertheless, I retain an aesthetic appreciation for the music, art, and architecture inspired over the centuries by religious faiths of many varieties. Moreover, it is impossible to ignore the fact that religion plays an important role in human affairs, mostly for good but too often for ill. It strikes me as more than a little odd that something as important as religion in human affairs has received relatively little attention from psychologists. As far as the emotions are concerned, I am particularly interested in how spiritual or mystical experiences, which are often cited as evidence for believing, might be explained in nonreligious terms (Averill, 1999a).

Without going into detail, I will outline briefly how a secular explanation might proceed. The first thing to note is that descriptions of mystical experiences are surprisingly uniform the world over (see, for example, the anthology

compiled by Aldus Huxley, 1985). The most commonly mentioned characteristic is a sense of oneness, a dissolution of boundaries that distinguish one thing from another and the self from all others. Second, the experience is described as highly meaningful, even life-changing. And, third, mystical experiences are associated with a sense of aliveness or vitality which, in the extreme, is often described as ecstasy or bliss.

The above three features (a sense of oneness, meaningfulness, and vitality) characterize what I will call classical or full-blown mystical experiences. Such experiences are extreme and hence, by definition, unusual. But so, too, are full-blown anxiety attacks or bouts of deep depression. And like anxiety and depression, mystical experiences can range in intensity from the mild to very intense; mild experiences are commonplace, and even moderately intense experiences are surprisingly frequent (Laski, 1968).

How might such experiences be explained? The common denominator, I suggest, is a controlled breakdown or deconstruction of the cognitive schemas by which we differentiate objects in the world. The result of such a breakdown would be the oft described “oceanic” feeling of oneness that is the hallmark of a mystical experience. However, in a world full of dangers, like the one in which humans evolved, such a condition could not last without fatal consequences. Cognitive schemas would have to be reconstructed quickly, and a new and perhaps more meaningful world created: hence, the sense of insight and greater knowledge frequently reported by mystics. And what about the experience of vitality, even ecstasy? An explanation of this feature is best approached indirectly. In everyday life, an inability to distinguish threatening from benign events would be stressful, to say the least. And, in fact, periods of anxiety and depression are commonly reported preludes to a mystical experience, what the Christian mystic, John of the Cross, described as “the dark night of the soul”. However, with adequate preparation, and in familiar surroundings, the experience of oneness and a new sense of meaning can be life-affirming, even ecstatic.

And how does the concept of emotions as

transitional social roles fit into this picture? It doesn’t, at least not without some qualifications. If a mystical experience involves a controlled deconstruction of cognitive schemas, it follows that the schemas that help constitute emotional roles would also be rendered inapplicable. However, the deconstruction of schemas is seldom complete. If the experience is to be blissful and not terrifying, a residual cognitive structure must remain (usually a religious or social ideology) that lends the experience some meaning. Otherwise, the experience would be like a psychotic breakdown.

Although drugs are seldom reported as triggers for mystical experiences (as opposed, say, to beauties of nature or sexual intimacies) drug-induced states are nevertheless suggestive: novice drug users often find their first experience disappointing, if not frightening; they must learn to be “high”, often under the tutelage of more experienced users. The same could be said of would-be mystics, who may undertake strenuous spiritual exercises in order to achieve their goal. Often, especially for lay persons, the “training” is more implicit than explicit. In a national survey, Greeley (1974) found listening to music to be the most frequently mentioned trigger for a mild mystical experience. I assume that most people do not listen to music as a training ground for mysticism, but for persons who have undergone extensive musical training, a mild or even intense mystical experience may occasionally be an unexpected byproduct.

A common criticism of your theory of emotions is that it neglects the biological dimension of emotions. What is the proper role of biology in a social constructionist account of emotions? Do you consider the basic emotion perspective to be irreconcilable with your own?

Social constructionism does not deny the importance of biology. Some components of an emotional syndrome may be largely innate (e.g., certain facial expressions), and some neural structures (e.g., the amygdala) may play a greater role in emotional than intellectual behavior. More broadly, biological systems of behavior — what used to be called *instincts*, such as aggression, sexual attraction, attachment, and flight from danger — may singly or in combination

contribute to the formation of some emotions (e.g., sexual attraction and attachment in the case of romantic love). However, no one-to-one relation exists between biological systems of behavior, either singly or in combination, and specific emotions (not even those that used to be called “basic”).

The “basic emotion perspective” is not necessarily irreconcilable with my own, especially as the former has evolved over the years. I do, however, have an argument with earlier versions, especially the tendency to link presumably basic emotions with certain facial expressions. What do facial expressions actually “express”? That is a question worth asking, but the answer should not rest on any a priori assumptions about the nature of emotions. Having said that, I believe that the basic emotion perspective has led to valuable research on the importance of facial expressions as an adjunct to spoken language, whether emotional or not.

As I alluded to in response to an earlier question, language provides a good analogy for understanding the role of biology in a social-constructionist account of emotions. Language is one of the most important biological adaptations of the human species. The full extent of the neurological correlates of language remains a mystery; however, the localization of some part-functions have long been known, for example, Broca’s area for the production of speech and Wernicke’s area for the perception of speech. But if you want to account for the difference, say, between English and Japanese, even the most sophisticated neurological assessment is unlikely to be of much aid. Rather, you must look to historical and comparative linguistics, and to the societies in which the languages are spoken today. And so it is with regard to emotional syndromes. If you want to understand the difference between anger in Western cultures and, say, *ikari* among the Japanese, it is best to look at historical and social factors and not neurobiology.

In short, a social-constructionist approach does not diminish the importance of understanding the biological and physiological bases of either language or emotion, especially when it comes to disorders of each.

In response to the difficulty of providing an

all-encompassing theory of emotions applicable to ordinary folk notions like anger, fear, shame, disgust, and so on, some theorists have doubted that ordinary emotion terms are proper objects of scientific investigation. What is your view on the matter?

The notion of “proper objects for scientific investigation” can be interpreted in two ways that are sometimes conflated. (a) Can folk concepts be used as *uneliminable* variables (i.e., indispensable constructs) in a comprehensive theory of emotion? My answer to this question is No. (b) Are folk concepts of emotion *relevant* to the development of a theory of emotion? My answer to this question is Yes.

Let me illustrate the reasons for these answers with an historical analogy, namely the theory of evolution originated by Darwin. With regard to point (a) although Darwin’s theory concerns the origin of species, no biological species enters the theory as an indispensable theoretical construct. This is because, as explained by Hull (1976), species are “logical individuals”. This concept requires some explication.

The most familiar example of an individual is, of course, a specific human being. You and I are individuals, born at a specific time and place and destined to die at another, and there will never be another *you* or *me* (even as a clone or in some alternative universe). Viewed as a unit, a species such as *Tyrannosaurus Rex* can also be considered an individual, that is, the species came into existence at one time and place and went extinct at another; and outside of a trick of genetic engineering, another species will not *evolve* that is exactly like *Tyrannosaurus Rex*. Historical events, such as the burial of Pompey by the eruption of mount Vesuvius, can also be considered individuals from a logical point of view.

Contrast the concept of a logical individual with a law of nature. A law, if it is a law and not simply an empirical generalization, must be spatiotemporally unrestricted, at least within its domain of applicability. For example, Newton’s law of gravity is not restricted to one particular time, nor to one location (although it may be superseded by an even more general law, such as Einstein’s, which has a greater range of applicability). It follows that no law of nature can

contain an uneliminable reference to logical individuals, for that would limit its applicability to a particular time and place. So much for point (a) mentioned above.

Let me turn now to point (b), namely, how species, although not *uneliminable* variables in a theory, may nevertheless be *relevant* to a theory. Prior to Darwin, biological species were generally viewed as “natural kinds”; that is, it was believed that a species could be defined in terms of essential features, much like a mathematical construct. As long as species were so conceived, the evolution of species was inconceivable. One species, it was reasoned, could no more evolve into another than a circle, say, could evolve into a square.

One of the major innovations of Darwin was to offer an alternative conception of species, namely, as sets of interbreeding individuals (Mayer, 1972). Since individual members of a species vary one from one another, species — being nothing more than aggregates of individuals — can evolve depending on which individuals reproduce more than others (assuming some degree of heritability).

As mentioned earlier, all constructionist approaches reject essentialism and thereby open the possibilities for change. But the lesson I want to draw from present analogy is somewhat different. If Darwin had not accepted species as theoretically relevant entities, but not as indispensable theoretical constructs, he would not have developed a theory to explain their origins by natural selection. Analogously, I suggest that if we do not take everyday emotions seriously, we will never develop a theory to explain their origins — and the possibilities for their change. Which brings me to your next question...

You have written a lot on emotional creativity: which emotions help creativity and why? Are there emotions that hinder creativity as well? Is there a way of becoming more creative by working on one's emotions, and if so how?

Persons tend to be more creative, regardless of domain (art, science, or whatever), when in a positive mood. Research supports what may seem intuitively evident: when in a positive mood thinking becomes more fluid, broader associations may be made among ideas, and the possibility of failure seems remote.

Unfortunately, things are seldom as straightforward as this description might suggest. For some persons, or for some phases of the creative process, a negative mood can also be helpful. The potential positive influence of negative states is illustrated most dramatically by the fact that creative writers and artists (but not scientists) are more prone to clinical depression than is the general population. But even that may not be an anomaly: among persons suffering from clinical depression, creative episodes tend to occur as depression lifts and the person enters a mildly manic phase.

My concern, however, is not with emotions that may facilitate or hinder creativity, but with emotional syndromes themselves as creative products. The possibility, almost inevitability, of emotional creativity is a straight-forward consequence of a social-constructionist view of emotion. What societies construct, individuals can reconstruct. If the reconstruction is not only novel, but also authentic (reflective of the individual's core values and interests) and effective (adaptive for the individual or group), we may speak of it as creative. If it fails to meet one or more of these criteria (novelty, authenticity, effectiveness), which, incidentally, are applicable to creativity in any domain, the emotional response may be regarded as eccentric or, worse, neurotic (Averill & Nunley, 1992).

Like most ideas, “emotional creativity” has ample precedents. Otto Rank (1932), a student of the arts and onetime disciple of Freud, believed that many neurotic syndromes reflect creative impulses that are expressed in ways detrimental to the individual. Starting from a different perspective, and focusing on the positive end of the neurotic-healthy spectrum, Abraham Maslow (1971) defined “primary” creativity as the ability to be inspired, to become totally immersed in the matter-at-hand, and to experience those “peak” moments that are “a diluted, more secular, more frequent version of the mystical experience” (p. 62). I would only add that emotional creativity is not limited to a few extreme (peak or mystical) experiences; a “spiritualization of the passions” (Nietzsche's phrase) can apply to a wide variety of emotions experienced in everyday life (Averill, 2009).

If emotional creativity is implicit in a social-constructionist perspective, why has it not been

more generally recognized, both in theory and in practice? One reason is that it is difficult to elicit an emotionally creative response in the laboratory, and hence to study its causal relations. I have therefore taken a different approach, namely, exploring the correlates of people who differ in emotional creativity as a trait. For this purpose, my students and I have constructed a 30-item Emotional Creativity Inventory (ECI). I want to acknowledge Carol Thomas-Knowles and Jenny Gutbezahl, in particular, for their assistance on this project.

The results of studies using the ECI have been presented elsewhere (Averill, 1999b). I will offer only a brief qualitative summary here. People who score high on the ECI, in comparison with their low scoring counterparts, tend to be more agreeable and open to experience, including mystical experiences in the sense discussed earlier, and they are better able to profit from solitude. High scorers are also better able to express their emotions creatively in images (drawings and collages) and in writing. These latter finding might reflect greater artistic or verbal talents on the part of high scorers, although attempts were made to control for such possibilities.

What emotional creativity is *not* is as important as what it is. For example, scores on the ECI are unrelated to general intelligence as measured by SAT scores. Emotional creativity can also be distinguished from emotional intelligence on both conceptual and empirical grounds (Averill, 2004, 2007). Finally, emotionally creative individuals are not especially reactive, that is, prone to respond with frequent or intense emotional outbursts.

The last point deserves brief elaboration because the stereotype of an emotionally creative persons might be someone who is flamboyant or who constantly seeks excitement and adventure. Of course, the stereotype may sometimes fit the case, but it is not necessarily the norm. Art, literature, and especially poetry may be better expressions of emotional creativity than is, say, sky diving or climbing a sheer mountain cliff (Oatley, 1999; Sundararajan & Averill, 2007).

Wordsworth (1805/1952) famously described poetry as “the spontaneous overflow of powerful feelings: it takes its origin from emotion recollected in tranquility” (p. 84). The operative

word here is *tranquility*. Inspired by Wordsworth, the philosopher John Stuart Mill (1833/1981) took the issue a step further. Mill considered poetry to be a way of educating and expanding the emotions; in other words, as a way of being emotionally creative. In this respect, he drew a distinction between the poet and people who “are perpetually engaged in hunting for excitement from without, [the latter] are invariably those who do not possess, either in the vigor of their intellectual powers or in the depth of their sensibilities, that which would enable them to find ample excitement nearer home.” That is an over generalization, clearly, but let’s follow Mill’s reasoning, for it leads to some relevant conclusions concerning emotional creativity.

Most literature (e.g., prose, drama, and rhetoric, no matter how eloquent) may also afford emotional excitement, Mill conceded, but also “of the kind that comes from without.” Poetry is different: Its object is “to paint the human soul truly.” “Great poets,” Mill asserted, “are often proverbially ignorant of life. What they know has come by observation of themselves: “they have found within them one highly delicate and sensitive specimen of human nature, on which the laws of emotion are written in large characters.” It follows, Mill concluded, that poetry “is the natural fruit of solitude and meditation,” not of active engagement in external affairs.

What should we conclude from this brief excursion on the relation between poetry and emotional reactivity? Not that all poetry involves emotional creativity (see Cupchik, 2016, Ch. 9, for a much more nuanced approach to the relation between emotions and poetry). And certainly the implication is not that we should all strive to become poets. Mill was no poet, although he was one of the major 19th century British philosophers. In fact, poetry is not even the issue. Rather, the issue is the source of a person’s “excitement” (to borrow Mill’s term). Emotional creativity presumes a rich inner life, and a willingness to explore and learn from it. The manner and extent that inner life gets expressed depends on an individual’s talents and circumstances.

Please list five articles or books that have had a deep influence on your thinking.

Wittgenstein, L (1953). *Philosophical Investigations* (G.E.M. Anscombe, Trans.). New York: Macmillan.

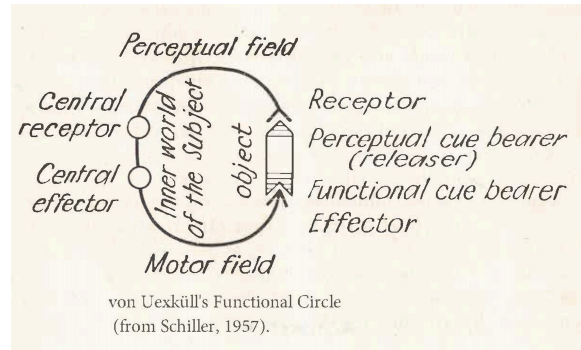
It is sometimes quipped that Wittgenstein's goal was to dissolve problems, not to solve them. This quip may contain a grain of truth, but it is also misleading. Many of the problems that plague us are due to what might be called word jams, that is, the piling up of words, usually taken out of context. Like log jams that impede the flow of a river, word jams impede the flow of thought. And as with log jams, sometimes dissolution is a necessary step toward a solution. Be that as it may, Wittgenstein's suggestions for dissolution do not come easy. The *Investigations* consist of a series of loosely connected paragraphs that place considerable demands on the reader.

Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson & Company

This is a felicitous complement to Wittgenstein's *Investigations*. Both Wittgenstein and Ryle are classed among a group of "ordinary language" or "analytic" philosophers, although the extent to which they influenced each other, if at all, is unclear. *The Concept of Mind* is an extended critique of Cartesian dualism, which Ryle dubs the legend of the "Ghost [mind] in the Machine [body]". To exorcize the ghost, Ryle replaces talk of mental events with references to behavioral dispositions (although he denies being a "behaviorist"). He draws on the emotions for many ghostly examples.

von Uexküll, J. (1957). *A stroll through the worlds of animals and men*. In C. H. Schiller (Trans. & Ed.), *Instinctive behavior: The development of a modern concept* (pp. 5-80). New York: International Universities Press. (Original work published 1934).

Would you like to know how the world appears to a tick? von Uexküll shows you that, with illustrations by G. Kriszat. And not just ticks, but animals of many kinds. Theoretically, one of von Uexküll's major contribution is the concept of a "functional circle", in which the perceived object is not only the bearer of stimulus properties ("releasers"), but also the bearer of cues that suggest a response, or what Gibson (1979) would later call "affordances." Changes in the internal state of the organism following an initial response



may add another dimension, as the functional circle spirals toward a goal.

Ford, C. S., & Beach, F. A. (1951). *Patterns of sexual behavior*. New York: Harper & Brothers.

To me, this is a model of how the study of emotion should proceed: choose a topic of interest and bring to bear on it as many sources of information as might be relevant. My only regret is that the book has not been updated by a third party, incorporating new information acquired over the past sixty years, especially with regard to possible biological and social determinants of gender identity. But that is a minor quibble; it is unlikely that additional data would alter the basic thrust of Ford and Beach's analysis, nor greatly modify the many examples they present of cultural variations in sexual behavior.

Berger, P. L., & T. Luckmann (1966). *The social construction of reality*. New York: Doubleday.

Written in almost telegraphic style, the authors bring the sociology of knowledge (cf. Scheler, Mannheim, and others) down to earth. Berger & Luckmann define "reality" as anything "that we recognize as having a being independent of our own volition"; that is, as anything we cannot "wish away". Although they give scant attention to the emotions per se, their definition of reality definitely includes the passions as traditionally conceived.

What are the most pressing questions emotion theorists should try to answer?

The integration of various perspectives (biological, psychological, social) is clearly a pressing issue. For as long as I can remember, people have been saying that, of course, all

behavior is a function of *both* heredity and environment, but little progress has been made beyond reiterating that truism. Recently, however, a subtle change in emphasis is taking place; according to Henrich (1916), evidence from a variety of disciplines suggest that culture has been the driving force behind the biological evolution of *Homo sapiens*. In other words, human beings as a species may be a social construction (see, also, Laland, 2017).

Another pressing need is to investigate a broad range of emotions as they function in everyday life. In the past, our theories have been built on too narrow a base — only a few emotions out of the hundreds recognized in everyday English, not to mention the many more recognized in other languages and cultures. I am encouraged by the variety of emotions currently being investigated, not just those considered biologically basic or socially important. Comparisons between closely related states, such as anger and annoyance, or loving and liking, I find particularly informative.

Which emotion, or set of related emotions, might yield theoretically significant insights is difficult to predict; so, it is important to keep an open mind. Before Darwin landed on the Galapagos Islands, who would have predicted that such mundane birds as finches would play a pivotal role in the development of a theory of evolution? So, too, may commonplace emotions yield unexpected theoretical insights.

What are you working on these days?

Since retirement, I have been working on having fun. In part, this involves sleeping late, traveling, and reading novels, mostly mysteries. More important is participation in a “Learning in Retirement” program. Currently roughly 300 people participate in the program — mostly former faculty from the five colleges in the area, and other retired professionals (teachers, physicians, engineers, lawyers, business people, artists). For each ten-week “semester” (Fall and Spring) some members propose seminar topics. Anyone who signs up for a topic is expected to prepare a presentation and lead a discussion. These are not occasions for “passive learning” (e.g., listening to a lecture). I usually sign up for two seminars a semester, one in the sciences and one in the humanities. That, together with



Jim's retirement dinner, with wife Judy and daughters Laurie and Andrea.

occasionally consulting with students and reviewing manuscripts, keeps me busy. Of course, not all is fun. Increasingly, I waste time looking for misplaced items, such as my car keys and cell phone.

References

- Arnold, B. (1960). *Emotion and personality*. (2 vols.). New York: Columbia University Press.
- Averill, J. R. (1968). Grief: Its nature and significance. *Psychological Bulletin*, 70, 721-748.
- Averill, J. R. (1969). Autonomic response patterns during sadness and mirth. *Psychophysiology*, 5, 399-414.
- Averill, J. R. (1973). Personal control over aversive stimuli and its relationship to stress. *Psychological Bulletin*, 80, 286-303. Named a "citation classic." (1983). *Current Contents: Social and Behavioral Sciences*, 15(21), 20.
- Averill, J.R. (1974). An analysis of psychophysiological symbolism and its influence on theories of emotion. *Journal for the Theory of Social Behavior*, 4, 147-190. (Reprinted in R. Harré & W. G. Parrott (Eds.) (1996). *The emotions: Social, cultural and biological dimensions* (pp. 204-228). London: Sage.)
- Averill, J.R. (1979). The functions of grief. In C. Izard (Ed.), *Emotions in personality and psychopathology*. New York: Plenum.
- Averill, J. R. (1980a). A constructivist view of emotion. In R. Plutchik and H. Kellerman (Eds.), *Emotion: Theory, research and experience: Vol. I. Theories of emotion* (pp. 305-339). New York: Academic Press. (Reprinted in A. G. Halberstadt & S.

- L. Ellyson (Eds.) (1990). *Social psychology readings: A century of research* (pp. 143-156). New York: McGraw-Hill.)
- Averill, J. R. (1980b). On the paucity of positive emotions. In K.R. Blankstein, P. Pliner, & J. Polivy (Eds.), *Assessment and modification of emotional behavior* (pp. 7-45). New York: Plenum Press.
- Averill, J. R. (1982). *Anger and aggression: An essay on emotion* New York: Springer-Verlag.
- Averill, J. R. (1985). The social construction of emotion: With special reference to love. In K. Gergen & K. Davis (Eds.), *The social construction of the person* (pp. 89-109). New York: Springer-Verlag.
- Averill, J. R. (1989). Stress as fact and artifact: An inquiry into the social origins and functions of some stress reactions. In C. D. Spielberger, I. G. Sarason, & J. Strelau (Eds.), *Stress and anxiety* (Vol. 12, pp. 15-38). Washington, DC: Hemisphere.
- Averill, J. R. (1990). Inner feelings, works of the flesh, the beast within, diseases of the mind, driving force, and putting on a show: Six metaphors of emotion and their theoretical extensions. In D. E. Leary (Ed.), *Metaphors in the history of psychology* (pp. 104-132). New York: Cambridge University Press.
- Averill, J. R. (1991). Emotions as episodic dispositions, cognitive schemas, and transitory social roles: Steps toward an integrated theory of emotion. In D. Ozer, J. M. Healy, Jr., & A. J. Stewart (Eds.), *Perspectives in personality* (Vol. 3a, pp. 139-167). London: Jessica Kingsley Publishers.
- Averill, J. R. (1999a). Spirituality: From the mundane to the meaningful — and back. *Journal of Theoretical and Philosophical Psychology*, 18, 101-126.
- Averill, J. R. (1999b). Individual differences in emotional creativity: Structure and correlates. *Journal of Personality*, 67, 331-371.
- Averill, J. R. (2004). A tale of two snarks: Emotional intelligence and emotional creativity compared. *Psychological Inquiry*, 15, 228-233.
- Averill, J. R. (2007). Together again: Emotion and intelligence reconciled. In G. Matthews, M. Zeidner, & R. D. Roberts (Eds.), *Emotional intelligence: Knowns and unknowns* (pp. 49-71). New York: Oxford University Press.
- Averill, J. R. (2009). Emotional creativity: Toward “spiritualizing the passions.” In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (2nd ed., pp. 249-257). New York: Oxford University Press.
- Averill, J. R., & Nunley, E. P. (1992). *Voyages of the heart: Living an emotionally creative life*. New York: The Free Press. (German edition: *Die Entdeckung der Gefühle: Ursprung und Entwicklung unserer Emotionen* (A. Pott, Trans.). Hamburg: Ernst Kabel Verlag, 1993.)
- Bowlby, J. (1961). The process of mourning. *International Journal of psycho-analysis*, 42, 317-340.
- Cupchik, G. C. (2016). *The aesthetics of emotion*. Cambridge, UK: Cambridge University Press.
- Dewey, J. (1895). The theory of emotion. II. The significance of emotions. *Psychological Review*, 2, 13-32.
- Dixon, T. (2003). *From passions to emotions: The creation of a secular psychological category*. Cambridge, England: Cambridge University Press.
- Durkheim, E. (1915). *The elementary forms of religious life*. (J. W. Swain, Trans.). New York: Macmillan.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Greeley, A. M. (1974). *Ecstasy: A way of knowing*. Englewood Cliffs, NJ: Prentice Hall.
- Hamburg, D.A. (1963). Emotions in perspective of human evolution. In P. H. Knapp (Ed.), *Expression of emotions in man*. New York: International Universities Press.
- Henrich, H. (2016). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University press.
- Hitchens, C. (2007). *The portable atheist*. Philadelphia: Da Capo Press.
- Hull, D. L. (1976). Are species really individuals? *Systematic Zoology*, 25, 174-191.
- Huxley, A. (1985). *The perennial philosophy*. London: Triad Grafton.
- James, W. (1884). What is an emotion? *Mind*, 9, 188-205.
- Laland, K. N. (2017). *Darwin’s unfinished symphony: How culture made the human mind*. Princeton University Press.
- Lange, C. G. (1922). *The emotions: A psychophysiological study*. In K. Dunlap

- (Ed.), *The Emotions* (pp. 33-90). New York: Hafner Publishing Co. (Original work published 1885)
- Laski, M. (1968). *Ecstasy: A study of some secular and religious experiences*. New York: Greenwood Press.
- Lazarus, R.S. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.
- Lazarus, R. S., Averill, J. R., Opton, E. M., Jr. (1970). Towards a cognitive theory of emotion. In M. B. Arnold (Ed.), *Feeling and emotion: The Loyola Symposium* (pp. 207-232). New York: Academic Press. (Reprinted in L. Levi (Ed.), *Society, stress, and disease*. Oxford University Press, 1971; N. Birsbaumer (Ed.), *Neuropsychologie der Angst*. München: Urban & Schwarzenberg, 1973; and G. Kahle (Ed.), *Theorie der Gefühle*. Frankfurt: Suhrkam Verlag, 1980.)
- Maslow, A. H. (1971). *The farther reaches of human nature*. New York: Viking Press.
- Mayer, E. (1972). The nature of the Darwinian revolution. *Science*, 176, 981-989.
- Mill, J. S. (1981). Thoughts on poetry and its varieties. In J. M. Robson & J. Stillinger (Eds.), *Collected Works of John Stuart Mill*, Vol. 1. *Autobiographical and Literary Essays* (pp.343-365). Toronto, Canada: University of Toronto Press. (Original work published 1833)
- Oatley, K. (1999). Fiction as cognitive and emotional simulation. *Review of General Psychology*, 3, 101-117.
- Parker, I. (Ed.) (1998). *Social constructionism, discourse, and realism*. London: Sage.
- Rank, O. (1932). *Art and artist*. (L. Lewison, Trans.). New York: Agathon Press.
- Sartre, J.P. (1948). *The emotions: Outline of a theory*. (B. Frechtman, trans.). New York: Philosophical Library.
- Solomon, R.C. (1976). *The passions*. Garden City, New York: Doubleday Anchor.
- Sundararajan, L., & Averill, J. R. (2007). Creativity in the everyday: Culture, self, and emotions. In R. Richards (Ed.), *Everyday creativity and new views of human nature* (195-220). Washington, DC: American Psychological Association.
- Wenger, M. A. (1950). Emotions as visceral action: An extension of Lange's theory. In M. L. Reymert (Ed.), *Feelings and emotions: The Mooseheart-Chicago Symposium* (pp. 3-10). New York: McGraw-Hill.
- Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Basil Blackwell & Mott.
- Wordsworth, W. (1952). Preface to second edition of lyrical ballads. In B. Ghiselin (Ed.), *The creative process* (pp. 83-84). Berkeley: University of California Press. (Original work published 1805)

ISRE Spotlight

Four Unwarranted Assumptions about the Role of Emotion in Moral Judgment

Nina Strohminger

Department of Psychology
Yale University
nina.strohminger@yale.edu

Pinning down the precise role emotion plays in moral judgment is as close to a white whale as moral psychology gets. Individually considered, emotional and moral processing are highly complex and amorphous phenomena. They resist precise measurement, or even a universally agreed upon definition. Our understanding of how these two are intertwined also depends critically on the logic undergirding our empirical methodology. There is, at present, an uncomfortably wide gap between what the standard methods have the power to show and the inferences we wish to draw from them. Yet these limitations have been largely ignored. In the absence of better tools, and in our zeal to answer the field's most pressing questions, it is easy to lose sight of what these methods can actually tell us.

There is, at present, no consensus on the extent or kind of influence emotions¹ have on moral judgment. At one end of the spectrum are those who deny that emotion impacts moral judgment at all. Sure, emotions *accompany* moral judgment, but their presence could be epiphenomenal. For instance, emotion could serve to motivate subsequent behavior, whilst remaining inert in the production of moral judgment (Mikhail, 2007; Huebner, Dwyer, & Hauser, 2009).

At the other end of the spectrum is the view

that emotion is deeply insinuated in moral judgment, so much so that it can lend moral resonance to otherwise benign actions (Haidt, Koller, & Dias, 1993; Nichols, 2004; Prinz, 2007). Hypnotizing subjects to experience disgust while reading stories causes them to think the characters are 'up to no good', even when the characters are engaged in innocuous activities like grocery shopping (Wheatley & Haidt, 2005). Adding emotional salience to conventional violations (like slurping one's soup) can make them seem like moral ones (like spitting into someone else's; Nichols, 2002). Of course, what counts as a 'moral' issue shape-shifts across communities and generations. Masturbation no longer seems the moral blight it once did; conversely, the moral significance Americans now attach to dietary choices would have been unimaginable a century ago. Some have proposed that emotion plays a role in this redistricting of the moral domain (Rozin, 1999; Horberg, Oveis, Keltner, & Cohen, 2009). Emotion may even be necessary for learning what is moral in the first place (Cushman, 2013). For instance, Robert Blair (1995) has argued that a deficit to the fear response in childhood leads to psychopathy in adulthood.

There are a variety of weaker versions of this claim. Perhaps emotion does not determine the boundaries of the moral domain, but intensifies disapprobation for an act we already believe is wrong (Pizarro, Inbar, & Helion, 2011; Huebner, 2015). Or perhaps emotion can only make acts seem morally wrong when they meet certain criteria, for instance if they concern principles of autonomy, community, and sanctity (Shweder, Much, Mahapatra, & Park, 1997; Graham, Haidt, & Nosek, 2009; Gray, Young, & Waytz, 2012). In short, emotion could nudge an opinion this way or that, without fundamentally altering moral perception.

Psychologists have not often drawn a distinction between the strong (constitutive) and weak (nudging) hypotheses (cf. Pizarro et al., 2011; Huebner, 2015). While this distinction is

¹ Researchers distinguish between emotions and other affective states, such as moods. However, when it comes to the processes underlying moral judgment, the critical question is whether *any* kind of affective state influences moral judgment. Therefore I treat

emotion and mood as interchangeable for the purposes of this discussion, as they both speak to the same problem.

important, both hypotheses boil down to a form of sentimentalism—the view that emotion plays a causal role in moral judgment. By and large, the methods discussed below have been used to advance some version of sentimentalism.

Despite how this evidence has usually been interpreted, these methods actually provide very little certainty about the role of emotion in moral judgment. Not only do these methods not allow us to adjudicate between the different flavors of folk sentimentalism, they don't count as terribly good evidence that emotion plays *any* kind of unique role in the production of moral judgment. The following is intended as a corrective.

One way of approaching this question is to measure how the propensity to experience emotion (trait emotion) correlates with moral beliefs. Individual differences in trait emotion can then be treated as a proxy for how emotion operates as a transient state for all members of the population. People who are high in trait disgust are less likely to give utilitarian moral judgments (Choe & Min, 2011) and more likely to endorse socially conservative viewpoints, such as that abortion and gay marriage are wrong (Inbar, Pizarro, & Bloom, 2009; Inbar, Pizarro, Iyer, & Haidt, 2012). Such findings could be used to argue that this is how disgust works as a transient emotional state as well: Anyone who renders a moral judgment while disgusted will become a little more conservative, a little less utilitarian. This basic maneuver—using trait emotion to infer the workings of state emotion—has been put forth for a variety of emotions, ranging from anger to empathy (Choe & Min, 2011).

It would certainly be convenient if trait emotion and state emotion were interchangeable; that is, if they reflected the same process, and led to the same output. However, we have very little evidence that the two produce comparable effects in the context of moral cognition. The most-studied emotion in this regard is disgust. And while putting people into a disgusted state has been shown to increase disapprobation for moral violations (Wheatley & Haidt, 2005; Schnall, Haidt, Clore, & Jordan, 2008), attempts to link disgust sensitivity to moral severity have largely failed (Nichols, 2002; Inbar, Pizarro, Knobe, & Bloom, 2009; Inbar, Pizarro, & Bloom, 2009;



Professor Nina Strohminger

Laakasuo, Sundvall, Drosinou, 2017, Pizarro, personal communication; Strohminger, unpublished data; cf. Jones & Fitness, 2008).

Outside of the moral psychology literature, similar inconsistencies appear. Dispositional sadness leads to less accurate and overgeneralized memory (Williams & Scott, 1988; Brittlebank, Scott, Williams, & Ferrier, 1993), whereas transient sadness leads to higher recall and greater detail-orientation (Storbeck & Clore, 2005; Forgas, Goldenberg, & Unkelbach, 2009; Forgas, 2010). Sadness does not have a uniform effect on cognitive performance. Rather, its effect depends on whether the individual is momentarily gloomy or perpetually so.

Unwarranted Assumption 1

(State/trait conflation)

The tendency to experience an emotion (trait) produces comparable effects to the temporary experience of that emotion (state).

We should probably not be surprised that stable personality traits do not always produce the same effect as fleeting emotional states. After all, the two represent quite different mental constructs. Trait emotion scales typically ask how frequently an emotion is experienced, or how intensely a stimulus elicits that emotion (Mehrabian & Epstein, 1972; Spielberger, Jacobs, Russell, & Crane, 1983; Haidt,

McCauley, & Rozin, 1994; Watson & Clark, 1994; Tangney, Dearing, Wagner, & Gramzow, 2000). A person who scores high on such a scale may do so because they have a low threshold for experiencing the emotion, or because of unique circumstances that increase emotional activation. So, while trait emotion is typically thought of as *dispositional*, it can also indicate chronic situational factors. To complicate matters further, the exact way that situational factors impact trait emotion appears to vary radically from emotion to emotion. Regular exposure to depressing life stimuli increases trait sadness (Spence, Najman, Bor, O’Callaghan, & Williams, 2002), whereas regular exposure to disgust-eliciting stimuli, such as cadavers and open wounds, lowers disgust sensitivity (Oaten, Stevenson, & Case, 2009). Whatever the underlying cause, a tendency to experience an emotion reflects persistent use of the neural pathways utilized in that emotion, which may lead to extensive and longstanding changes to how the emotion is processed. In short, what we call “trait emotion” bears no straightforward relationship to state emotion, and may not itself be a unified phenomenon.

There is another problem with emotion trait measures, one so obvious I almost forgot to mention it. Most studies that use emotion trait scales use them in correlational, not experimental, designs. This means that cause cannot be disentangled from effect. This limitation applies to another popular method in the field, which I’ll call *motus operandi*. *Motus operandi* involves presenting stimuli (in this case, moral scenarios), then measuring the emotions they evoke. This emotional response can be measured in a variety of ways, from self-report (Russell & Giner-Sorolla, 2011; Lee & Ellsworth, in press) to behavior (Cohen, Nisbett, Bowdle, & Schwarz, 1996) to psychophysiology (Chapman, Kim, Susskind, & Anderson, 2009) to neural activation (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Moll et al., 2002). *Motus operandi*, when it leads to a measurable emotional response, is often used as evidence that moral judgment involves emotion. For instance, moral dilemmas involving a ‘personal’ deontological moral violation (e.g. requiring an agent to push a person off of a footbridge in order to save five lives) generate greater activation in the brain areas associated with emotion (Greene, Nystrom,

Engell, Darley, & Cohen, 2004). Greene and colleagues have used these results to argue that deontological reasoning relies on emotions more heavily than utilitarian reasoning.

But there is nothing in *motus operandi* that shows emotion has perturbed moral judgment in the slightest. All this method reveals is that emotion arises at approximately the same time as moral processing. The observed emotional response could be epiphenomenal; indeed, it could occur after the moral judgment has been formed. Even neuroimaging work, which has played such a pivotal role in reigniting discussions of emotion’s role in moral judgment, remains agnostic on this point. Imagine that *motus operandi* had higher temporal resolution, and could demonstrate that emotion preceded moral judgment—this still would not eliminate the possibility that the emotional response was epiphenomenal. (That said, to the extent that we do have mental chronometry evidence, it reveals that emotion is not the first morally relevant information to be processed by the brain; Decety & Cacioppo, 2012.) A few researchers have noted this problem with *motus operandi*, and have offered a variety of alternative explanations: perhaps emotion draws attention to morally relevant information (Decety, Michalska, & Kinzler, 2012), intensifies moral salience (Pizarro et al., 2011), or serves to motivate appropriate moral action (Huebner et al., 2009).

Unwarranted Assumption 2

(Temporal correlation/causation conflation)

Appearance of emotion at approximately the same time as moral judgment shows that the emotion plays a causal role in moral judgment.

To get purchase on the causality question, researchers must introduce a manipulation. When we talk about how emotion influences moral judgment, we generally mean how emotion, *as elicited by the moral stimulus*, helps give rise to the final judgment. This is referred to as *integral* emotion. A graphic description of a rape may incite a strong emotional reaction that influences our disapprobation of the act. But emotions can also linger, influencing subsequent, unrelated judgments. Having eaten a bad egg salad sandwich before entering the courtroom may have an undue influence on our verdict—this is *incidental* emotion. The fact that

Unwarranted Assumptions

emotions are susceptible to misattribution is a useful quirk of the cognitive system which can be exploited by researchers.

It is generally treated as a matter of course that incidental emotion works the same as integral emotion. Indeed, this assumption forms the basis for the prominent use of emotion priming paradigms in the moral psychology literature. In an emotion priming paradigm, researchers induce an emotion in subjects—for instance, by showing a video or a series of pictures. Subjects then complete an ostensibly unrelated task, such as rating a series of acts for moral permissibility, hoping that the emotion in the first part of the experiment will spill over into the second part. The emotion priming paradigm relies on several assumptions, actually, but the one we will focus on at the moment is that misattributed (incidental) emotion works the same way as correctly attributed (integral) emotion.²

There is some, limited, evidence that incidental and integral emotion work in comparable ways. For instance, adding disgusting details to conventional violations moralizes them (integral emotion; Nichols, 2002), much as hypnosis-induced disgust moralizes the behavior of fictional characters (incidental emotion; Wheatley & Haidt, 2005). But it is far from an established certainty, and surprisingly, the issue has not been the object of systematic study.

There is good reason, though, to think that the two may not be equivalent. By definition, integral emotion can arise only after morally relevant information is first presented. When an emotion is primed, however, it is active before (often well before) any moral information is shown. Emotion can act in an orienting capacity, making certain information more salient (Phelps, Ling, & Carrasco, 2006; Harrison, Skau, Franconeri, Lu, & Chang, 2013). This is significant, because moral judgment is highly susceptible to the order in which information is presented (Petrinovich & O'Neill, 1996; Young & Saxe, 2008; Lombrozo,

2009; Schwitzgebel & Cushman, 2012; Wiegmann, Okan, Nagel, 2012). For instance, whether children weight intent or outcome more heavily in a moral scenario depends on which is presented closer to the time of the final judgment (a recency effect of sorts; Feldman, Klosson, Parsons, Rholes, & Ruble, 1976; Parsons, Ruble, Klosson, Feldman, & Rholes, 1976; Surber, 1982). Because emotion influences informational salience, its timing alongside the details of a moral scenario may differentially affect how those details are processed. It would therefore be wise to exercise caution when considering whether the effects of incidental emotion generalize to integral emotion.

Unwarranted Assumption 3

(Integral/incidental conflation)

Inducing an emotion and observing its effect on subsequent, unrelated tasks is equivalent to how the emotion works when it is integral to the task.

Increasingly, emotion priming paradigms have been recruited to make a very specific kind of claim: that observing an effect of incidental emotion on moral judgment demonstrates a special link between the two processes. Incidental emotion, so the argument goes, will only influence a subsequent cognitive process when emotion is usually a part of that process. Valdesolo and DeSteno (2006) found that a positive mood induction (in the form of a humorous video clip) affected judgments on a moral dilemma that appeals to deontological reasoning but not one that appeals more to utilitarian intuitions. This finding has been used to argue that deontological judgments rely on emotion but utilitarian judgments do not. Similar claims have been made for incidental emotion across various types of moral evaluation (Wheatley & Haidt, 2005; Strohminger, Lewis, & Meyer, 2011; Inbar, Pizarro, & Bloom, 2012; Seidel & Prinz, 2012; Cummins & Cummins, 2012).

The problem with this reasoning is emotion affects all manner of decidedly unemotional

² Emotion priming paradigms can also be used to motivate a purely pragmatic argument: This is what happens when an unrelated emotion is introduced to a sterile field. Results like this become important when one is trying to design environments that allow

judgment and decision-making to proceed in a fair and unbiased manner. For this more industrial use, I have no quarrel: but it speaks nothing to cognitive process.

cognitive processes. Indeed, one would be hard-pressed to find a mental operation impenetrable to incidental emotion; the list includes attention (Phelps et al., 2006; Stefanucci, Proffitt, Clore, & Parekh, 2008; Sherman, Haidt, & Clore, 2012), memory (Bower, 1981; Levine & Pizarro, 2004), test performance (Wine, 1971), math (Bryan & Bryan, 1991), assessments of monetary value (Lerner, Small, & Loewenstein, 2004), and heuristic reasoning (Kassam, Koslov, & Mendes, 2009; Inbar & Gilovich, 2011). Suppose we were to discover that putting people into a sad mood makes them perform worse on a math exam. No one would use this to argue that emotion is required for mathematical cognition. Yet this is precisely how emotion priming studies have been used to understand moral cognition.

Besides, one could easily make the opposite argument, that incidental emotion should only pervade cognitive processes that *don't* use emotion, because those do not already have that channel occupied by competing emotional processing. In effect, emotion priming could be treated as an affective version of a cognitive load task (Payne, Hall, Cameron, & Bishara, 2010), where incidental emotion is most effective when the mind is not already engaged with other emotional processes. I have never seen anyone advocate for this view, though the logic strikes me as no less sound.

Unwarranted Assumption 4

(Emotion manipulation as normal operation)
If emotion influences a subsequent, unrelated task, then emotion is usually recruited to that task.

In a more sophisticated version of this argument, selective use of certain emotions reflects a functional relationship between emotion and moral domain. If an incidental emotion, such as disgust, affects a moral judgment, this could be used to argue that disgust is a 'moral' emotion (Haidt, 2003). But this is a questionable strategy, as it would quickly lead one to rope in every emotion as a moral emotion: happiness, sadness, mirth, lust (Ariely & Loewenstein, 2006; Strohminger et al., 2011). Few emotions are active exclusively within the moral domain.

Or imagine that you have mapped individual emotions onto specific moral problems—disgust

is recruited for violations of purity, anger for violations of autonomy, and contempt for violations of community (Rozin, Lowery, Imada, & Haidt, 1999). In that case, you might expect to find that incidental disgust selectively affects judgments related to purity. However, no clear pattern using emotion priming has ever emerged aligning specific emotions to particular kinds of moral content. Studies show that disgust can selectively impact judgments of purity violations (Horberg et al., 2009), that disgust selectively does *not* impact purity violations, even though it impacts other types of moral judgment (Ugazio, Lamm, & Singer, 2012), that disgust impacts some purity violations but not others (Inbar, Pizarro, & Bloom, 2012), that disgust has no selective effect across moral domains (Schnall et al., 2008), and that disgust impacts moral judgments on scenarios where there is no moral content present (Wheatley & Haidt, 2005). I do not think we should be shocked by the inadequacy of emotion priming to resolve such subtleties. If we can find that incidental disgust affects enjoyment of paintings and cartoons (Strohminger, 2013), why should it be any more selective when it comes to moral judgment? It may well be that certain emotions are more often recruited for certain types of moral problem, but using incidental emotion to adjudicate on this issue is folly.

Quite recently, some replications and meta-analyses have come out suggesting that incidental disgust does not impact moral judgment at all (Landy & Goodwin, 2015; Johnson et al., 2016). At first blush, this may seem like a terrible blow—and indeed it is a blow, but only for the existence of the basic effect. This analysis is not at all devastating for the deeper question of the role disgust normally plays in moral judgment, for the reasons outlined above. Incidental disgust tells us very little—perhaps as little as nothing at all—about whether disgust is normally recruited during moral cognition, and if so, how disgust impacts moral judgment.

Some may take comfort in this. But this should also be an opportunity for sober reflection. Emotion priming was long held as a critical piece of evidence as to the constitutive role of disgust in moral judgment, when all along it was just a tinker toy. This fundamental mistake

marks an opportunity for us to begin thinking more carefully about our methods. The kinds of claims they can be used to support, and the kinds of claims they cannot.

References

- Ariely, D., & Loewenstein, G. (2006). The heat of the moment: The effect of sexual arousal on sexual decision making. *Journal of Behavioral Decision Making*, 19, 87–98.
- Blair, R. (1995). A cognitive developmental approach to morality: Investigating the psychopath. *Cognition*, 57, 1–29.
- Bower, G. (1981). Mood and memory. *American Psychologist*, 36, 129–148.
- Brittlebank, A., Scott, J., Williams, J. M., & Ferrier, I. (1993). Autobiographical memory in depression: State or trait marker? *The British Journal of Psychiatry*, 162, 118–121.
- Bryan, T., & Bryan, J. (1991). Positive mood and math performance. *Journal of Learning Disabilities*, 24, 490–494.
- Chapman, H., Kim, D., Susskind, J., & Anderson, A. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, 323, 1222–1226.
- Choe, S. Y., & Min, K.-H. (2011). Who makes utilitarian judgments? the influences of emotions on utilitarian judgments. *Judgment and Decision Making*, 6, 580–592.
- Cohen, D., Nisbett, R. E., Bowdle, B. F., & Schwarz, N. (1996). Insult, aggression, and the southern culture of honor: An “experimental ethnography”. *Journal of Personality and Social Psychology*, 70, 945–960.
- Cummins, D. D., & Cummins, R. C. (2012). Emotion and deliberative reasoning in moral judgment. *Frontiers in Psychology*, 3, 393–408.
- Cushman, F. (2013). Action, outcome and value: A dual-system framework for morality. *Personality and Social Psychology Review*, 17, 273–292. (Unpublished manuscript.)
- Decety, J., & Cacioppo, S. (2012). The speed of morality: a high-density electrical neuroimaging study. *Journal of Neurophysiology*, 108, 3068–3072.
- Decety, J., Michalska, K. J., & Kinzler, K. D. (2012). The contribution of emotion and cognition to moral sensitivity: a neurodevelopmental study. *Cerebral Cortex*, 22, 209–220.
- Feldman, N. S., Klosson, E. C., Parsons, J. E., Rholes, W. S., & Ruble, D. N. (1976). Order of information presentation and children’s moral judgments. *Child Development*, 556–559.
- Forgas, J. P. (2010). Affect and global versus local processing: The processing benefits of negative affect for memory, judgments, and behavior. *Psychological Inquiry*, 21, 216–224.
- Forgas, J. P., Goldenberg, L., & Unkelbach, C. (2009). Can bad weather improve your memory? an unobtrusive field study of natural mood effects on real-life memory. *Journal of Experimental Social Psychology*, 45, 254–257.
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96, 1029–1046.
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23, 101–124.
- Greene, J., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44, 389.
- Greene, J., Sommerville, R., Nystrom, L., Darley, J., & Cohen, J. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293, 2105–2108.
- Haidt, J. (2003). The moral emotions. In R. J. Davidson, K. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 852–870). New York, NY: Oxford University Press.
- Haidt, J., Koller, S., & Dias, M. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, 65, 613–628.
- Haidt, J., McCauley, C., & Rozin, P. (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences*, 16, 701–713.
- Harrison, L., Skau, D., Franconeri, S., Lu, A., & Chang, R. (2013). Influencing visual judgment through affective priming. In

- Proceedings of the 2013 annual conference on human factors in computing systems (chi).*
- Horberg, E., Oveis, C., Keltner, D., & Cohen, A. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology*, 97, 963–976. doi: 10.1037/a0017423
- Huebner, B. (2015). Do emotions play a constitutive role in moral cognition? *Topoi*, 34, 427–440.
- Huebner, B., Dwyer, S., & Hauser, M. (2009). The role of emotion in moral psychology. *Trends in Cognitive Sciences*, 13, 1–6.
- Inbar, Y., & Gilovich, T. (2011). Angry (or disgusted), but adjusting? the effect of specific emotions on adjustment from self-generated anchors. *Social Psychological and Personality Science*, 2, 563–569.
- Inbar, Y., Pizarro, D., & Bloom, P. (2009). Conservatives are more easily disgusted than liberals. *Cognition and Emotion*, 23, 714–725.
- Inbar, Y., Pizarro, D., & Bloom, P. (2012). Disgusting smells cause decreased liking of gay men. *Emotion*, 12, 23–27.
- Inbar, Y., Pizarro, D., Iyer, R., & Haidt, J. (2012). Disgust sensitivity, political conservatism, and voting. *Social Psychological and Personality Science*, 3, 537–544.
- Inbar, Y., Pizarro, D., Knobe, J., & Bloom, P. (2009). Disgust sensitivity predicts intuitive disapproval of gays. *Emotion*, 9, 435–439.
- Johnson, D. J., Wortman, J., Cheung, F., Hein, M., Lucas, R. E., Donnellan, M. B., . . . Narr, R. K. (2016). The effects of disgust on moral judgments: Testing moderators. *Social Psychological and Personality Science*, 7, 640–647.
- Jones, A., & Fitness, J. (2008). Moral hypervigilance: The influence of disgust sensitivity in the moral domain. *Emotion*, 8, 613–627.
- Kassam, K. S., Koslov, K., & Mendes, W. B. (2009). Decisions under distress: Stress profiles influence anchoring and adjustment. *Psychological Science*, 20, 1394–1399.
- Laakasuo, M., Sundvall, J., & Drosinou, M. (2017). Individual differences in moral disgust do not predict utilitarian judgments, sexual and pathogen disgust do. *Scientific Reports*, 7, 45526.
- Landy, J. F., & Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science*, 10, 518–536.
- Lee, S. W. S., & Ellsworth, P. C. (in press). Maggots and morals: Physical disgust is to fear as moral disgust is to anger. In K. Scherer & J. Fontaine (Eds.), *Components of emotional meaning: A sourcebook*. Oxford, UK: Oxford University Press.
- Lerner, J., Small, D., & Loewenstein, G. (2004). Heart strings and purse strings: Carryover effects of emotions on economic decisions. *Psychological Science*, 15, 337–341.
- Levine, L., & Pizarro, D. (2004). Emotion and memory research: A grumpy overview. *Social Cognition*, 22, 530–554.
- Lombrozo, T. (2009). The role of moral commitments in moral judgment. *Cognitive Science*, 33, 273–286.
- Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, 40, 525–543.
- Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in cognitive sciences*, 11, 143–152.
- Moll, J., de Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Mourão-Miranda, J., Andreiuolo, P. A., & Pessoa, L. (2002). The neural correlates of moral sensitivity: A functional magnetic resonance imaging investigation of basic and moral emotions. *The Journal of Neuroscience*, 22, 2730–2736.
- Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition*, 84, 221–236.
- Nichols, S. (2004). *Sentimental rules: On the natural foundations of moral judgment*. New York, NY: Oxford University Press.
- Oaten, M., Stevenson, R., & Case, T. (2009). Disgust as a disease-avoidance mechanism. *Psychological Bulletin*, 135, 303.
- Parsons, J. E., Ruble, D. N., Klosson, E. C., Feldman, N. S., & Rholes, W. S. (1976). Order effects on children's moral and achievement judgments. *Developmental Psychology*, 12, 357–358.
- Payne, B. K., Hall, D. L., Cameron, C. D., & Bishara, A. J. (2010). A process model of

- affect misattribution. *Personality and Social Psychology Bulletin*, 36, 1397–1408.
- Petrinovich, L., & O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. *Ethology and Sociobiology*, 17, 145–171.
- Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*, 17, 292–299.
- Pizarro, D., Inbar, Y., & Helion, C. (2011). On disgust and moral judgment. *Emotion Review*, 3, 267–268.
- Prinz, J. (2007). *The emotional construction of morals*. Oxford, UK: Oxford University Press.
- Rozin, P. (1999). The process of moralization. *Psychological Science*, 10, 218–221.
- Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, 76, 574–586.
- Russell, P. S., & Giner-Sorolla, R. (2011). Moral anger, but not moral disgust, responds to intentionality. *Emotion*, 11, 233.
- Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, 34, 1096–1109. doi: 10.1177/0146167208317771
- Schwitzgebel, E., & Cushman, F. (2012). Expertise in moral reasoning? order effects on moral judgment in professional philosophers and non-philosophers. *Mind & Language*, 27, 135–153.
- Seidel, A., & Prinz, J. (2012). Mad and glad: Musically induced emotions have divergent impact on morals. *Motivation and Emotion*, 1–9.
- Sherman, G. D., Haidt, J., & Clore, G. L. (2012). The faintest speck of dirt disgust enhances the detection of impurity. *Psychological science*, 23, 1506–1514.
- Shweder, R., Much, N., Mahapatra, M., & Park, L. (1997). The “big three” of morality (autonomy, community, divinity) and the “big three” explanations of suffering. In A. Brandt & P. Rozin (Eds.), *Morality and health: Interdisciplinary perspectives* (pp. 119–169). New York, NY: Routledge.
- Spence, S. H., Najman, J. M., Bor, W., O'Callaghan, M. J., & Williams, G. M. (2002). Maternal anxiety and depression, poverty and marital relationship factors during early childhood as predictors of anxiety and depressive symptoms in adolescence. *Journal of Child Psychology and Psychiatry*, 43, 457–469.
- Spielberger, C. D., Jacobs, G., Russell, S., & Crane, R. S. (1983). Assessment of anger: The state-trait anger scale. In C. D. S. J. N. Butcher (Ed.), *Advances in personality assessment* (Vol. 2, pp. 159–187). Hillsdale, NJ: Lawrence Erlbaum.
- Stefanucci, J. K., Proffitt, D. R., Clore, G. L., & Parekh, N. (2008). Skating down a steeper slope: Fear influences the perception of geographical slant. *Perception*, 37, 321.
- Storbeck, J., & Clore, G. L. (2005). With sadness comes accuracy; with happiness, false memory mood and the false memory effect. *Psychological Science*, 16, 785–791.
- Strohming, N. (2013). *The hedonics of disgust* (Unpublished doctoral dissertation). The University of Michigan.
- Strohming, N., Lewis, R., & Meyer, D. (2011). Divergent effects of different positive emotions on moral judgment. *Cognition*, 119, 295–300.
- Surber, C. F. (1982). Separable effects of motives, consequences, and presentation order on children's moral judgments. *Developmental Psychology*, 18, 257–266.
- Tangney, J. P., Dearing, R. L., Wagner, P. E., & Gramzow, R. (2000). The test of self-conscious affect-3 (tosca-3) [Computer software manual]. Fairfax, VA.
- Ugazio, G., Lamm, C., & Singer, T. (2012). The role of emotion for moral judgments depends on the type of emotion and moral scenario. *Emotion*, 12, 579–590.
- Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychological Science*, 17, 476–477.
- Watson, D., & Clark, L. A. (1994). The PANAS-X: Manual for the positive and negative affect schedule - expanded form [Computer

- software manual]. Iowa City, IA.
- Wheatley, T., & Haidt, J. D. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science, 16*, 780–784. doi: 10.1111/j.1467-9280.2005.01614.x
- Wiegmann, A., Okan, Y., & Nagel, J. (2012). Order effects in moral judgment. *Philosophical Psychology, 25*, 813–836.
- Williams, J., & Scott, J. (1988). Autobiographical memory in depression. *Psychological Medicine, 18*, 689–695.
- Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76*, 92–104.
- Young, L., & Saxe, R. (2008). The neural basis of belief encoding and integration in moral judgment. *Neuroimage, 40*, 1912–1920.

Feature Article: Empathy

Recent Work in the Philosophy of Empathy

Derek Matravers

Philosophy Department
Open University
derek.matravers@open.ac.uk

It is an exciting time in Anglo-American philosophy in the study of empathy. In addition to a host of articles, there have recently been two significant collections of essays published (Coplan & Goldie, 2011; Maibom, 2014). There have been two books – one a serious monograph, and the other an overview (Stueber 2006, Matravers, 2017). Finally, a true sign of the debate having arrived, a handbook has just appeared, which contains no fewer than 33 papers (Maibom, 2017). However, it would be a mistake to imagine that there is a consensus, or, indeed, even an agreement on the nature (or existence) of the phenomenon under discussion. In this contribution, I will briefly outline the recent history of empathy within philosophy, and then say something about what I take to be the most interesting question currently engaging the field.

The absence of agreement is not much of a surprise, given the history of the concept. Aspects of the contemporary discussion bear a marked similarity to what certain Enlightenment thinkers, most particularly Adam Smith and David Hume, called ‘sympathy’. Unfortunately, one such similarity is a failure precisely to demarcate the phenomenon; to distinguish it (or not) from such phenomena as emotional contagion or sympathy (in the modern sense). The term (‘empathy’) has a much more recent history. As is well known, it was coined by Edward Titchener in 1909, as a translation of the German, ‘Einfühlung’. The debate from which this German neologism sprang was also confused; the core seemed to be that beauty is a matter of our ‘animating’ parts of the world by projecting our inner states into them. This notion of ‘projection’ was subject to excoriating criticism by Vernon Lee in 1910 (Lee

& Anstruther-Thomson, 1912), in 1914 the theory was magisterially dismissed by E.F. Carritt: ‘we have here nothing but an attempt to explain in figurative language an unconscious process by which some beautiful objects may have become so’ (Carritt, 1962, pp. 191). Discussion then largely disappears until around the 1980s, when two separate areas of philosophy become hotly debated. The first was the debate as to how we know the contents of another’s mind – I shall refer to this as the ‘mind reading’ debate. The second was the debate over the nature of the emotions. I shall start by examining this distinction, and then move to discussing empathy and the emotions – as that will be of most interest to readers of this newsletter.

The mindreading debate sprung out of a dissatisfaction with the consensus view – functionalism. Functionalism holds that person A works out what person B is thinking by relying on some general theory that connects perceptual inputs and behavioural outputs with our mental states. That is, A is supposed to rely on some tacitly known theory which is full of theorems of the sort ‘In circumstance C, if X desires Y, and believes that Z-ing is the best way to get Y, then X will Z’. Such a view faced a number of problems both in principle and in practice. Independently, Jane Heal (in the UK) and Robert Gordon (in the US) proposed an alternative (Gordon, 1986; Heal, 1986). Instead of using a theory to discover what someone else is thinking, we (broadly) imagine that we in that person’s situation, note what we would think in such circumstances, and attribute that thought to the person. This notion of taking someone else’s perspective drew on thinking on empathy. Indeed, for a while the view was known as ‘the empathy view’, before settling on ‘simulation theory’ (Davies & Stone, 1995, pp. 1).

Possibly stemming from the same basic dissatisfaction with an excessive focus on functionalism, philosophical work on the emotions also became popular in the 1980s. For a while empathy merely had a walk-on part, with few if any mentions. This changed when Peter Goldie, in his 2000 monograph, *The Emotions*, discussed it more extensively (Goldie, 2000). Such discussions linked with the discussion of simulation theory, and the touch-paper was lit.

However, it is not obvious the two debates have much in common. The first, mindreading debate, was basically about working out what other people are thinking and the second, emotion debate, about feeling what other people are feeling.

Those engaged in the mindreading debate did not deny that some people used the term ‘empathy’ in a way that linked it to affective states. Their view was, rather, that this was not where the concept could do most work; a view they (rightly) linked to the tradition out of which the term had emerged – the discussion of *Einfühlung*. Hence, Kirsten Stueber defines ‘empathy’ as ‘a form of *inner or mental imitation for the purpose of gaining knowledge of other minds*’ (Stueber, 2006, pp. 28). Indeed the gulf between the mindreading debate and the emotions debate soon got wider. Working out what people will think – that is, working out the inferential connections they will make – has greater prospects of success than the rather messy business of working out what people will feel. Hence, the project of ‘gaining knowledge of other minds’ has sometimes left the emotions behind altogether, and been restricted to purely cognitive states (Heal, 1988). Across the divide, work on empathy as an affective state has side-lined the cognitive. In the emotion debate, this definition of ‘empathy’ by Heidi Maibom would be fairly standard: ‘S empathizes with O’s experience of emotion E in C if S feels E for O as a result of: believing or perceiving that O feels E, or imagining being in C’ (Maibom, 2014, pp. 3).

Having laid out the contrast, I will, as indicated above, focus specifically on a problem with empathetic emotions. Philosophers are interested in understanding emotional empathy both because it is an interesting phenomenon in itself and because of the role such emotions play in other areas of enquiry. Particularly fruitful is the link to morality. Hume and Smith, mentioned above, both made empathy (or, as they called it ‘sympathy’) foundational for their moral theories (see in particular Hume (1739-40) and Smith (2002)). However, as contemporary debate shows, understanding the role empathy plays in our moral thinking and conduct is particularly difficult. This debate is not purely (or even principally) conducted by philosophers. Among the psychologists, Martin Hoffman holds views



Professor Derek Matravers

‘empathy as the bedrock of morality’ (Hoffman 2011, pp. 96), while Daniel C. Batson holds that ‘empathy-induced altruism is neither moral nor immoral; it is amoral’ (Batson, 2014, pp. 47). A similar division can be found in philosophy. Michael Slote thinks empathy is the key to morality (Slote, 2007) while Jesse Prinz thinks that ‘empathy is not necessary for the capacities that make up basic moral competence’ and ‘can interfere with the ends of morality’ (Prinz, 2011, pp. 211-213). The philosophical view is that much of this debate (which, through Paul Bloom has become part of popular culture (Bloom, 2016)) is greatly helped by the kind of distinction-drawing, conceptual clarity, and interpretation of evidence, in which the discipline excels.

On, then, to the empathetic emotions. Let us restrict ourselves to the usual case where someone imagines themselves occupying the perspective of another (I shall ignore various problems with this, some of which I will return to below). I shall call the person who is being empathetic ‘the empathiser’ and the person for whom empathy is felt ‘the target’. Consider an example in which the target feels sad because their dog, Fido, has died; that is, the emotion (E)

is the target's sadness at Fido. The question is: what does the empathiser feel?

According to Maibom's definition, the empathiser feels E. That, however, is a complicated claim that needs some unpacking. One issue is how finely we want to individuate E. Let's be old-fashioned cognitivist about this, and individuate emotions by their cognitive content: that is, E is 'sadness at the death of Fido'. Could this be what the empathiser feels?

One might think not for at least two reasons. First, the relation between the empathiser and the death of Fido is unlikely to be the same as the relation between the target and the death of Fido (Matravers, 2011). This is, the latter is likely to have ramifications throughout the target's mental economy; a sense of loss, starting every time a dog barks or a door is pushed open, weeping in the pet-food isle, and so on. Imagining being in the target's situation, as we are assuming the empathiser does, is unlikely to bring about the same effects. One could reply that the empathiser feels what the target feels but not to the same degree. However, it is not clear that is coherent. Consider a parallel example. The target is in love with Susan; he cannot concentrate, his mind returns to her constantly, he is intoxicated by the sound of her voice, the smell of her hair. Clearly, the empathiser (who, let us assume, is not in love with Susan) does not feel all (or any) of that. It is not obvious it makes sense to say that he feels it, but just not to the same degree.

One solution is to say that the cause and object of the empathiser's emotion is not the death of Fido, but rather some object that plays an equivalent role in the empathiser's life (say, Rover – a dog of his that died). That is, in imagining being in C, the empathiser takes on the target's perspective, but substitutes Rover for Fido and thus generates the kind of ramifications in her (the empathiser's) mental economy as the death of Fido does in the target's mental economy. Something like this might be the case, but it needs a bit of finessing. At first pass it does not look to be empathy at all; it looks, rather, as if the target's situation has triggered the empathiser into their own emotion – feeling sad for themselves at the death of their own dog. That might, however, not be the best way of specifying the cause and object of the empathiser's emotion. Instead we might say this: essentially, the cause

and object of the empathiser's emotion is something of a sort that would be the appropriate cause and object of the target's emotion. That is, what the empathiser tries to do is to find something (anything) in her (the empathiser's) life that matches the role the death of Fido has in the target's life. So it is not really right to say that the empathiser is feeling sad about the death of Rover; they are feeling sad about some state of affairs whose nature is dictated by the object of the target's emotion. As a matter of fact, that thing is the death of Rover, but the most informative way of describing what the empathiser is feeling sad about is not 'sad-at-the-death-of-Rover' but 'sad-at-something-that-matches-what-the-target-feels-sad-about'.

Let us let that worry lurk in the background as I shall return to it below. For the second reason for holding that E is not 'sadness at the death of Fido' is the more interesting: that it cannot be the emotion the empathiser has in mind as she could have that emotion without implicating the target; but the target clearly is implicated.

The obvious way to sort this out is to have the target, rather than Fido, as the object of the empathiser's emotion. This is the literal reading of Maibom's definition above: S feels E for O. *Prima facie*, this is not an attractive solution. If the object of the empathiser's emotion was the target, rather than the dog, that would be sympathy (feeling sad for someone who is sad) rather than empathy (feeling the sadness of someone who is sad).

Maibom herself is aware of the problem and has this to say:

'Feeling for' is to be understood broadly so as to include cases where I am angry with a person because that person wronged you, where, in a sense, you are neither the object nor the subject of my emotion. What makes it a case of empathic anger is that I am feeling it not directly as a sort of objective moral anger, but rather I feel it on your behalf. (Maibom, 2014, pp. 5; see also Maibom, 2017)

What is it to feel an emotion on someone's behalf? Stephen Darwall discusses this under the name 'proto-sympathetic empathy'.

A person grieves the loss of his child, and in sharing his grief projectively my focus is on the child who was lost, not the person whose grief I share. When, however, I turn my attention to what it must be like to live with this loss, I focus on the person himself and the ways his grief pervades and affects his life. Before my thought was: What a terrible thing – a precious child is lost. Now my thought is: What a terrible thing for him – he has lost his precious child. (Darwall, 1997, pp. 271)

Darwall does not go into detail about how this might be done. Here is a possible scenario. The empathiser imaginatively identifies with the target. In doing so, he imagines that his child (that is, the target's child) has died. However, the empathiser also has a perspective on how, within the imaginative project, he (the empathiser) feels. Assuming that the attempt at empathy has been successful, this will also be how the target feels. Hence, the empathiser is able to note, and track, how the target feels. From this second-order perspective, the empathiser can (a) note that they are feeling 'on the target's behalf' and (b) form a judgment, on the basis of what they are feeling, about what the target is feeling. This would seem to implicate the target in the right kind of manner.

The field is split as to whether or not this scenario is plausible. Peter Goldie has been particularly sceptical. His scepticism is rooted in the possibility that the empathiser can replicate, in any reliable way, the flow of thoughts and emotions around the target's mental economy. In particular, the psychological role of emotional dispositions is not the kind of thing that can be duplicated in imagination (Goldie, 2011). This is slightly different to the problem I considered earlier: that the same object (the death of Fido) might have different effects on different people because of different people will have different relations to that object. Here the issue is more that the structure of the mind (in particular, the relation between our emotional dispositions and our emotional experiences) cannot be replicated in imagination. I shall not dwell on this here;

suffice to say, it might be an advantage if we are able to implicate the target without relying on this controversial model.

Let us focus on (b): what is the judgement? One option would be for the empathiser to decide on the sort of emotion they feel – and given what Darwall has said, that might be complicated – and judge that the target feels an emotion of that sort. However, they might instead use the emotion they are feeling as an exemplar; they simply make the judgement that the target feels like *this*; where 'this' refers to the emotion they (the empathiser) is feeling. This solution has recently been put forward by Kendall Walton (see also Joel Smith (Smith, 2015)).³ In Walton's example, Emily is the empathiser and Oscar is the target.

Emily's judgement or impression is not merely that '*I am panicked, and so is Oscar*,' but rather, '*Oscar is as I am, like this*.' She can appropriately say, 'I know how it is with him' or 'I know how he feels,' where 'know' carries a connotation of intimacy, acquaintance... Notice that the content of what she knows is in propositional form: *She knows that Oscar feels like this*. But this is propositional knowledge of a special kind, with the sample taking the place of a linguistic predicate in the formulation of what she knows. (Walton, 2015, pp. 9)

The key point to notice about Walton's view is that, at least with respect to (b), we do not have to rely on the controversial scenario detailed above. All that needs to be the case is that the empathiser is experiencing an emotion they believe is the one felt by the target. There is no reason to restrict the method by which they generated this emotion to imaginatively identifying with the target. They could, for example, simply recall what it was like for them when they were in the target's circumstances, and, having dredged up that emotion from memory, claim that the target feels *like this* (where 'this' is the emotion they are feeling). Walton's definition of empathy postulates no necessary link to the imagination: 'I propose to

³ Strictly, Walton does not think it is necessary for the empathiser to judge; they can merely experience

the target 'as feeling "like this"' (Walton, 2015, pp. 9). I shall ignore this complication.

define “empathy” as, simply, using some aspect of one’s current mental state as a sample to understand another person, in the way I have described, i.e., judging or experiencing the target person to be feeling “like this” (Walton, 2015, pp. 9-10).

Walton’s account not only has the advantage of not relying on the controversial scenario, it also avoids the earlier problem that, as the empathiser stands in a different relation to the object of the target’s emotion as does the target, the effect on the empathiser will be different. My suggested solution to that problem was that the empathiser focusses on ‘some object that plays an equivalent role in the empathiser’s life’. This fits perfectly into Walton’s proposal. Empathising with you on the death of your dog involves me recalling or imagining the death of my dog, thus creating an emotional experience, and judging that you feel like *this*.

Thus, if we focus just on (b), there is good reason to favour Walton’s proposal. What, however, about (a): that we are feeling ‘on the target’s behalf’? Well, this is partly a matter of judgement. On some views, empathy is the key to moral and spiritual progress – the process of identifying with another, taking on their issues, and thus being motivated to help them. Walton’s view certainly seems ‘thinner’ – perhaps too thin to give us everything a pre-theoretical understanding of empathy (if there is such a thing) promises. This would not bother Walton; he is a ‘theory builder’ in philosophy; rather than someone who analyses concepts as he finds them, he is happy to take them away, tidy them up, and return them to the linguistic wild. If we follow Walton, as I have suggested we should, it reveals that the link between empathy and the imagination might not be as tight as we might have assumed. The pay-off is a neat solution as to how we can feel our own emotion, implicate the target in that emotion, and yet not have to explain this in ways that some have found incredible.

Empathy is a good example of a topic where philosophy can make progress by working together with other disciplines including psychology (in its clinical, developmental, and social forms), phenomenology and hermeneutics, and cognitive neuroscience. Indeed, it is an area in which progress will only be secure once certain philosophical issues are bashed out: what exactly

is it to ‘take someone else’s perspective’? What is it to feel an emotion ‘on behalf of’ another? What is egoism, altruism, morality, and how are they related? What are the differences between emotional contagion, empathy, and sympathy, and when are they important? These, and other issues, leave plenty of work to do.

References

- Batson, C. D. (2014). Empathy-induced altruism and morality: No necessary connection. In H. L. Maibom (Ed.), *Empathy and Morality*. (pp. 41-58). Oxford, UK: Oxford University Press.
- Bloom, P. (2016). *Against Empathy: The Case for Rational Compassion*. New York, NY: Ecco.
- Carritt, E. F. (1962). *The Theory of Beauty*. London, UK: Methuen.
- Coplan, A. & Goldie, P. (2011). *Empathy: Philosophical and Psychological Perspectives*. Oxford, UK: Oxford University Press.
- Darwall, S. (1997). Empathy, sympathy, care. *Philosophical Studies* 89, 261-282.
- Davies, M. & Stone, T. (1995). Introduction. In M. Davies & T. Stone (Eds.), *Folk Psychology: The Theory of Mind Debate* (pp. 1-44). Oxford, UK: Blackwell.
- Goldie, P. (2000). *The Emotions*. Oxford, UK: Clarendon Press.
- Goldie, P. (2011). Anti-empathy. In A. Coplan & P. Goldie (Eds.), *Empathy: Philosophical and Psychological Perspectives* (pp. 302-317). Oxford, UK: Oxford University Press.
- Gordon, R. M. (1986). Folk psychology as simulation. In M. Davies & T. Stone (Eds.), *Folk Psychology: The Theory of Mind Debate* (pp. 60-73). Oxford, UK: Blackwell.
- Heal, J. (1986). Replication and functionalism. In *Mind, Reason, and Imagination* (pp. 11-27). Cambridge, UK: Cambridge University Press.
- Heal, J. (1988). Co-cognition and off-line simulation: Two ways of understanding the simulation approach. In *Mind, Reason and Imagination* (pp. 91-114). Cambridge, UK: Cambridge University Press.
- Hoffman, M. L. (2011). Empathy, justice, and the law. In A. Coplan and P. Goldie (Eds.), *Empathy: Philosophical and Psychological*

- Perspectives* (pp. 230-254). Oxford, UK: Oxford University Press.
- Hume, D. (1739-40). *A Treatise of Human Nature*. London, UK: Thomas Longman.
- Lee, V. & Anstruther-Thomson, C. (1912). *Beauty and Ugliness and Other Studies in Psychological Aesthetics*. London, UK: John Lane.
- Maibom, H. L. (2014). *Empathy and Morality*. Oxford, UK: Oxford University Press.
- Maibom, H. L. (2014). Introduction. In H. L. Maibom (Ed.), *Empathy and Morality* (pp. 1-40). Oxford, UK: Oxford University Press.
- Maibom, H. L. (2017). Affective empathy. In H. L. Maibom (Ed.), *The Routledge Handbook of Philosophy of Empathy* (pp. 22-32). London, UK: Routledge.
- Maibom, H. L., Ed. (2017). *The Routledge Handbook of Philosophy of Empathy*. London, UK: Routledge.
- Matravers, D. (2011). Empathy as a Route to Knowledge. In A. Coplan and P. Goldie (Eds.), *Empathy: Philosophical and Psychological Perspectives* (pp. 19-30). Oxford, UK: Oxford University Press.
- Matravers, D. (2017). *Empathy*. Cambridge, UK: Polity.
- Prinz, J. (2011). Is empathy necessary for morality? In A. Coplan and P. Goldie (Eds.), *Empathy: Philosophical and Psychological Perspectives* (pp. 211-229). Oxford, UK: Oxford University Press.
- Slote, M. (2007). *The Ethics of Care and Empathy*. London, UK: Routledge.
- Smith, A. (2002). *The Theory of Moral Sentiments*. Cambridge, UK: Cambridge University Press.
- Smith, J. (2015). What Is Empathy For? *Synthese*, 194, 709-722.
- Stueber, K. R. (2006). *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences*. Cambridge, Mass., MIT Press.
- Walton, K. (2015). Empathy, Imagination, and Phenomenal Concepts. In *In Other Shoes: Music, Metaphor, Empathy, Existence* (pp. 1-16). Oxford, UK: Oxford University Press.

Feature Article: Empathy

Empathy and its Development: What is Missing?

Alexandra Main

Psychological Sciences
University of California, Merced
amain@ucmerced.edu

Philosophers have discussed the phenomenon of empathy since the 1800s. Robert Vischer (1873) coined the term “*empathie*” (translated as “feeling into”), which was later translated into “empathy” by Theodor Lipps (Lipps, 1914). However, the empirical study of empathy in psychology did not gain much traction until the 1980s. Behaviorism had fallen out of fashion at this point, and psychology researchers became more interested in the “black box” of what was going on inside people’s heads. This spurred a burst of research on emotion. Though there has been debate over whether empathy is an “emotion” per se, one would be pressed to find an empathy researcher who did not believe that emotion is inherent to empathy.

How do psychologists study empathy?

In psychology, empathy is often defined as feeling what another feels. This seems simple enough, but upon further reflection, there are several caveats. Is a person empathizing if they feel what someone else is feeling because of emotional contagion (i.e., the tendency to “catch” others’ emotions), or is a cognitive understanding of why someone is feeling a certain way necessary? How are other emotional responses to others’ emotions, including sympathy and compassion, related to or distinct from empathy? How are vicarious emotional responses to others’ emotions that do not involve feeling the same emotion as another, such as feeling angry in response to someone being hurt, related to or distinct from empathy? Is empathy always a precursor to positive social behaviors, such as helping another person? What about when we feel too overwhelmed by someone else’s distress

(i.e., personal distress) to help? Are we empathizing in this case?

These questions have led many researchers in psychology to argue that empathy is a multidimensional phenomenon. Davis (1983) argued that empathy has both affective (i.e., feeling what someone else is feeling) and cognitive (i.e., understanding someone else’s perspective) components. In other words, it is not enough to simply feel sad when we see someone who is sad, but we must also appreciate that the person is sad because their dog died. This view continues to pervade much of the thinking in the field and most empathy researchers distinguish between affective and cognitive components of empathy (see Brown et al., 2017; Cuff et al., 2016). Research with adolescents suggests that these different components of empathy have distinct developmental timelines (e.g., Van Lissa et al., 2014; Davidov et al., 2013) and there may be distinct neurological processing for more cognitively mediated vs. “pure” forms of empathy (Zaki & Ochsner, 2012).

An excellent example of research on the multidimensional nature of empathy comes from a study by Zaki and colleagues (Zaki et al., 2009). In this study, research participants watched videos of individuals telling emotional stories while in an fMRI scanner. Participants were instructed to rate using a dial how positively or negatively the storyteller was feeling at that moment while they were talking about their emotional event. These ratings were then mapped onto the ratings the storytellers had made of their own feelings to develop a measure of empathic accuracy – how accurate the participants were at judging the storytellers’ emotions. The researchers found that certain systems in the brain associated with understanding others’ mental states were particularly active during this task; specifically, the superior temporal sulcus and medial prefrontal cortex (Zaki et al., 2009).

Other neuroscience research examining individuals’ responses to the pain of others activate more “primitive” neural networks, including the anterior cingulate cortex, brainstem, and cerebellum (Singer et al., 2004). The researchers argue that the processes involved in “mentalizing,” or trying to understand other people’s point of view, is a neurologically distinct process from the more visceral reaction we

experience when we witness someone experiencing an emotion, particularly distress or pain (Zaki & Ochsner, 2012). This suggests that not only are there multiple psychological ingredients that are necessary for empathy, but there is physical evidence in the brain of this multidimensionality as well.

What do we know about the development of empathy?

How we define and operationalize empathy has important implications for understanding its development. The work of Paul Bloom, Michael Tomasello, and others suggests that babies may not need to learn empathy; rather, they are capable of empathy from birth – entering the world as altruistic beings. In a clever early study, infants were either played a recording of another newborn's cry, a synthetically-produced cry, or silence (Sagi & Hoffman, 1976). Infants were significantly more likely to cry in response to another newborn's cry than in the other two conditions. Many have interpreted this to mean that infants have a rudimentary form of empathy from birth. Though there are competing explanations to this phenomenon, it is clear that even infants seek to connect with others, be it through the social smile emerging around 6-8 weeks (Anisfeld, 1982), intersubjectivity and joint attention emerging around 6 months (Butterworth & Cochran, 1980), or helping behavior emerging around 15 months of age (Zahn-Waxler et al., 1992).

Unlike research with adults that typically relies on self-report measures to study empathy, researchers needed to be a bit more creative in studying empathy in toddlers and children. Early studies with toddlers involved systematically observing children's responses to another's simulated distress (either the mother or an experimenter) by coding their facial expressions and behaviors (e.g., facial expressions of concern, personal distress, hypothesis testing, helping). This early work showed that between 12-24 months of age, children became more sophisticated in their responses to others' distress (Zahn-Waxler et al., 1992). For example, in response to a mothers' simulated distress (e.g., "hurting" herself by bumping her foot), infants engaged in more reparative behaviors as they got older, such as engaging in prosocial behavior



Professor Alexandra Main

(actively trying to relieve the parent's distress). Infants also become less egocentric as they get older. Specifically, in the same study, infants engaged in fewer self-focused responses to the parent's distress, including crying or rubbing the infant's own foot, as they got older.

While very young children are sensitive to others' distress, they are not necessarily very effective at using their empathic responses to motivate successful helping of others. For example, younger children may bring an experimenter or the caregiver their own toy in response to the adult's distress. This reflects an understanding that the person is in distress, but a failure to demonstrate the capacity to respond to the adult in a way that will reduce the adult's distress. Martin Hoffman's seminal writings in the field of psychology argued that such a progression from "egocentric empathy" to veridical empathy was fundamentally a cognitive process involving self-other distinctions that become more differentiated and complex with development (Hoffman, 1982).

Other researchers posited that such development of other-oriented empathy fundamentally involves the developing ability to regulate one's own distress in response to another's distress. We might initially feel upset when we see someone hurt, but if we can turn our attention away from our own distress and focus on the other person, we will be better equipped to help them. Nancy Eisenberg demonstrated that children's effortful control (a temperamentally-based individual differences in the ability to deploy attention and inhibit undesirable behaviors) is a precursor to sympathy. She argues

that sympathy is a well-regulated empathic response that involves care and concern for others, as opposed to personal distress, which involves poorly-regulated distress in response to another's distress that is focused on the self rather than on the other (Eisenberg et al., 2013). Though Eisenberg's and Hoffman's theories have their subtle distinctions, they overlap in their view of empathic development as fundamentally progressing from a "pure" emotional response to a more cognitively-controlled process.

What is missing in the study of empathy?

The research described thus far has provided great insights into the development and neurological substrates of empathy (see Uzefovsky & Knafo-Noam, 2017 and Zaki & Ochsner, 2012 for reviews). However, the focus of empathy research has traditionally emphasized the internal *experience* of the empathizer. This focus has often come at the detriment to understanding how empathy is an *interpersonal, relational process*. My colleagues and I have argued that current conceptualizations of empathy as feeling what another is feeling are not necessarily *wrong*, but they fail to capture the richness of the empathic process and do not provide many practical implications for how people can learn to be better empathizers (Main et al., 2017).

Consider the following example: A mother finds out that her teenage son has been sending text messages after he is supposed to be in bed. As punishment, she takes his phone away. The adolescent responds by withdrawing into his room (a common behavior among adolescents). The parent may interpret this withdrawal as anger at his phone being taken away. However, in reality, she failed to see that her son was in fact sad because he viewed his mother's punishment as a lack of trust.

This example highlights several elements of empathy that are often missing in empirical research. First, empathy is inherently interpersonal in nature. If we think of empathy only as something happening within the mother (i.e., the empathizer), she was empathizing with her son by putting herself in his shoes. However, she made an incorrect assumption about the motivation for her son's behavior. Second, empathy is contextually-bound. In this situation,

the adolescent withdrew from his mother, which made it more difficult for her to empathize. However, in another situation, the adolescent may have been more willing to disclose the reasons behind his feelings, which would have facilitated empathy between mother and son, likely leading to a more positive long-term relationship. Third, empathy is a dynamic process. The mother may have *initially* assumed her son was angry about his phone being taken away, but if we were to stop our assessment there, we would miss whether she persisted in this belief, or if she attempted to interact with her son to determine if her initial judgment was accurate. Without an appreciation of the corrective processes involved in real-time empathizing with others, we are missing the deeply dynamic nature of empathy. Fourth, this example highlights that the mother's curiosity is central to effectively empathizing with her son. Below I provide examples from the real world as well as from research to highlight these underemphasized aspects of empathy.

Empathy is interpersonal. Although it may seem obvious that empathy is an interpersonal process – with whom would we empathize if not another person? – empathy is typically studied in solitary contexts. This is likely based on convenience (it is easier to study one person than two!), but theoretical perspectives that view that empathy is an internal experience. However, this neglects the fact that empathy is an interactive social process that depends not only on the empathic tendencies of the empathizer, but also on the openness or resistance of the social partner to being empathized with (Hollan, 2008; Ickes et al., 1997).

Consider a follow-up study by Zaki and colleagues in which the researchers examined the perceivers' empathic accuracy and how empathic they generally reported themselves to be on a self-report questionnaire (Zaki, Bolger, & Ochsner, 2008). They found that people who scored highly on an empathy questionnaire, unsurprisingly, tended to be more accurate in their empathic judgments while observing the storytellers telling emotional stories. However, the perceivers were only accurate in the cases when the storyteller was highly expressive. This suggests that even highly empathic people may struggle to empathize if the person with whom

they are trying to empathize is emotionally blunted or closed-off.

When the person we are trying to empathize with gives us feedback, such as telling us we are wrong and displaying an angry facial expression suggesting we misunderstood them, this helps us better empathize with those around us. In other words, when we consider the openness or resistance of others to being empathized with, we can appreciate the fact that empathy is a two-way street.

Empathy is contextually-bound. Another limiting factor to much of the research on empathy is that it is typically studied as a static trait (you expressed empathy or not in this situation, you are an empathic person or not). However, empathy is highly context-dependent. Sharing someone else's emotional experience may be adaptive in some contexts, such as when we are motivated to help another by sharing their pain, but sharing someone else's emotions may not always be the most effective route to empathizing. For instance, empathy is a crucial piece of effective conflict resolution (e.g., Levenson & Ruef, 1992), but sharing someone else's negative emotions during a conflict is likely to lead to conflict escalation rather than resolution (Gottman et al., 2014).

John Gottman and his colleagues have conducted several decades of careful research in which they have observed couples discussing issues of conflict with one another and analyze the second-by-second emotion expression of each partner. The researchers consistently find that couples who allow their negativity to escalate, rather than engage in attempts to repair the relationship after one partner expresses negative emotion, report lower marital satisfaction and are far more likely to get divorced (see Driver et al., 2012). Thus, empathy in a conflict context may best be characterized as validating or showing interest in someone else's emotions, not matching their emotion (Main et al., 2017).

Empathy is a dynamic process. Empathy does not occur at a finite point in time, but rather unfolds dynamically over time. Despite this, much of the empirical research on empathy uses self-report questionnaires about dispositional tendencies to empathize or requires participants to make one-time judgments about others' emotions. However, empathy is often a corrective

process. This is evident when we use others' facial, verbal, and postural cues to determine whether we are accurate or not in our assessments of our partner's emotions, such as when a wife displays a flash of anger toward her husband who immediately tries to make her feel better before validating her sadness.

Some individuals may not pick up on such cues, but others may use these cues to "correct" their assumptions. Without an appreciation of these individual variations and how contextual factors might influence these differences, we may be missing some very fundamental aspects of the empathy process.

The role of curiosity in empathy. Much of the research on empathy involves whether individuals successfully label a person's emotions. However, empathy in the real world is more complex. Consider being at a party and you seeing a friend in the corner looking sad. It is not enough to identify your friend's expression as communicating sadness. Rather, truly empathizing involves understanding *why* your friend is sad, which requires a curiosity about her with her environment.

This issue is particularly relevant in clinical settings. Recent research in behavioral medicine has addressed the potential for psychological burnout among physicians who share their patients' distress (see Halpern, 2014). Thus, other empathic behaviors, such as curiosity about another's emotional point of view, may be more effective in situations involving caring for others. Indeed, curiosity about a patient's feelings and perspective facilitates patient disclosure, which predicts healthy outcomes in patients (Suchman et al., 1997). Simply labeling a patient as angry about a cancer diagnosis or displaying sympathy toward a patient who does not want others feeling sorry for her would not be considered empathic in real life, so it makes little sense that empathy is usually empirically studied in this way.

Empathy is culturally-situated. It is important to note that everything described thus far is culturally-situated. How we define empathy, express it to others, respond to others' emotions, and how much different cultures value empathy can vary considerably.

Douglas Hollan's ethnographic work highlights cultural differences in social expectations about how we express emotions and

respond to the emotions of others – both of which have tremendous implications for how we think about empathy in different cultures (Hollan, 2012). For example, it is considered culturally inappropriate among the Yapese (a rural cultural group in Indonesia) to appeal for sympathy from others, while in another cultural group, the Toraja, expressions of vulnerability are key to positive social interactions. Furthermore, while we typically define empathy in Western society as an internal state (though I hope I have convinced the reader that we need to move beyond this), other cultural groups view empathy as an active, instrumental response (e.g., exchanging goods) rather than a passive sharing of another's emotional experience (von Poser, 2011).

Taken together, understanding empathy from a cross-cultural necessitates that even the basic definition of empathy needs to be understood from a culturally-situated point of view.

Opportunities for future research

Psychologists have made tremendous progress toward better understanding of the internal experience of empathy. This emphasis on internal experience is likely because psychologists are inherently interested in what is going on in the human mind. But humans do not exist in a vacuum. In the real world, we are very rarely asked to label other people's emotions explicitly without any feedback and we do not often encounter situations where we have only a one-time opportunity to empathize with others. Often psychology researchers emphasize tight control over experimental designs in the quest for internal validity. However, this may jeopardize the external validity of much of this research (Campos et al., 2011).

Research from other fields, including anthropology, behavioral medicine, and linguistics, can serve to complement the tightly-controlled research emphasized in psychology. For example, conversation analysis, a qualitative technique used in linguistics research, allows researchers to track behaviors over time, including empathy. One study found that empathic displays (e.g., validating statements) occurring early in a conversation may shut down a conversation by discouraging one's partner from opening up, but such statements were highly

effective at promoting empathic understanding later in the conversation (Kupetz, 2014). An important goal for future research would be to incorporate and develop quantitative methodologies that balance internal and external validity to better capture the phenomenon of empathy.

Another opportunity for deepening our understanding of empathy would be to allow for more behavioral flexibility in our dependent measures. For example, while allowing subjects to interact with the target of empathy in a study adds a layer of complexity and messiness to a study, we may develop a deeper understanding of how people empathize with others in real time.

Empirical research on empathy typically starts with *a priori* assumptions of what kinds of behaviors are expected to be empathic (e.g., facial expressions of concern, overt helping behavior), but this reflects an emphasis on the *form* of empathy rather than its *function*. I have highlighted here that in some contexts affect matching might be appropriate, but in others, curiosity might be a more empathic response. Shifting our focus from assuming certain behaviors are inherently empathic to a greater appreciation of the contextual appropriateness of such behaviors would move the study of empathy toward greater clarity.

Despite the challenges associated with studying empathy from an interpersonal and ecologically valid perspective, it is time for empirical research to take a more dynamic and relational approach to shed light on how we can better empathize with others across the complex situations real life presents to us.

References

- Anisfeld, E. (1982). The onset of social smiling in preterm and full-term infants from two ethnic backgrounds. *Infant Behavior and Development*, 5, 387-395.
- Brown, M. M., Thibodeau, R. B., Pierucci, J. M., & Gilpin, A. T. (2017). Supporting the development of empathy: The role of theory of mind and fantasy orientation. *Social Development*.
- Butterworth, G., & Cochran, E. (1980). Towards a mechanism of joint visual attention in human infancy. *International Journal of Behavioral Development*, 3, 253-272.

- Campos, J. J., Walle, E. A., Dahl, A., & Main, A. (2011). Reconceptualizing emotion regulation. *Emotion Review*, 3, 26-35.
- Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review*, 8, 144-153.
- Davidov, M., Zahn-Waxler, C., Roth-Hanania, R., & Knafo, A. (2013). Concern for others in the first year of life: Theory, evidence, and avenues for research. *Child Development Perspectives*, 7, 126-131.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113-126.
- Driver, J., Tabares, A., Shapiro, A. F., & Gottman, J. M. (2012). Couple interaction in happy and unhappy marriages: Gottman Laboratory studies. In F. Walsh (Ed.), *Normal family processes: Growing diversity and complexity (4th ed.)*, (pp. 57-77). New York, NY, US: Guilford Press.
- Eisenberg, N., Betkowski, J., & Spinrad, T. L. (2013). Age-related changes in empathy-related responding. In D. Hermans (Ed.), *Changing emotions*, (pp. 17-23). New York, NY, US: Psychology Press.
- Halpern, J. (2014). From idealized clinical empathy to empathic communication in medical care. *Medicine, Health Care and Philosophy*, 17, 301-311.
- Hoffman, M. L. (1982) Affect and moral development. *New Directions for Child Development*, 16, 83-103.
- Hollan, D. W. (2008). Being there: On the imaginative aspects of understanding others and being understood. *Ethos*, 36, 475-489.
- Hollan, D. W. (2012). Emerging issues in the cross-cultural study of empathy. *Emotion Review*, 4, 70-78.
- Ickes, W. J., Marangoni, C., & Garcia, S. (1997). Studying empathic accuracy in a clinically relevant context. In W. J. Ickes (Ed.), *Empathic accuracy* (pp. 282-310). New York, NY, US: Guilford Press.
- Kupetz, M. (2014). Empathy displays as interactional achievements – Multimodal and sequential aspects. *Journal of Pragmatics*, 61, 4-34.
- Levenson, R. W., & Ruef, A. M. (1992). Empathy: A physiological substrate. *Journal of Personality and Social Psychology*, 63, 234-246.
- Lipps, T. (1914): *Asthetik: Psychologie des Schoenen und der Kunst, Vol. 1, Grundlegung der Ästhetik*, Leipzig/Hamburg: Voss.
- Main, A., Walle, E. A., Kho, C., & Halpern, J. (2017). The interpersonal functions of empathy: A relational perspective. *Emotion Review*.
- Sagi, A. & Hoffman, M. L. (1976). Empathic distress in the newborn. *Developmental Psychology*, 12, 175.
- Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech: The beginnings of interpersonal communication*. Cambridge: Cambridge University Press.
- Uzevovsky, F., & Knafo-Noam, A. (2017). Empathy development throughout the life span. In J. A. Sommerville & J. Decety (Eds.), *Social cognition: Development across the life span* (pp. 71-97). New York, NY, US: Routledge/Taylor & Francis Group.
- van Lissa, C. J.; Hawk, S. T., de Wied, M., Koot, H. M., van Lier, P., & Meeus, W. (2014). The longitudinal interplay of affective and cognitive empathy within and between adolescents and mothers. *Developmental Psychology*, 50, 1219-1225.
- Vischer, R. (1873): *Über das optische Formgefühl: Ein Beitrag zur Ästhetik*, Leipzig: Credner.
- Von Poser, A. (2011). Bosmun foodways: Emotional reasoning in a PNG life-world. In D. W. Hollan & C. J. Throop (Eds.), *The anthropology of empathy: Experiencing the lives of others in Pacific societies* (pp. 169–193). Oxford, UK: Berghahn.
- Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., Chapman, & M. (1992). *Developmental Psychology* 28, 126-136.
- Zaki, J., Bolger, N., & Ochsner, K. (2008). It takes two: The interpersonal nature of empathic accuracy. *Psychological Science*, 19, 399-404.

- Zaki, J., Weber, J., Bolger, N., & Ochsner, K. (2009). The neural bases of empathic accuracy. *PNAS*, *106*, 11382-11387.
- Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls, and promise. *Nature Neuroscience*, *15*, 675-680.

Feature Article: Empathy

The Therapeutic Effects of Empathy in Healthcare

Jodi Halpern

School of Public Health
University of California, Berkeley
jhalpern@berkeley.edu

Increasing attention is being paid to the role of empathy in healthcare, but too little is understood about which specific aspects of empathy are therapeutic and how those aspects work together in clinical encounters.

Traditionally, physicians believed that they could employ a special kind of “detached concern” in which they label patients’ emotions from the outside, looking in, but need not experience actual empathy in which they vividly imagine what patients are going through. They argued that this detachment was essential for them to be objective and not burn out.

Starting about 20 years ago, scholarly work began to emerge that argued against these assumptions and investigated the therapeutic value of emotional empathy (resonating with patients’ feelings and attuning to them non-verbally) as well as cognitive empathy (seeing things from the patients’ perspective as opposed to one’s own).

In *From Detached Concern to Empathy*, I argued for an integrated affective-cognitive model for clinical empathy in which affects guide what the physician is able to imagine about the patient’s experience (Halpern, 2001). This work was followed by empirical research (reviewed below) showing that the physician’s combined affective-cognitive engagement does increase the effectiveness of medical care.

This essay considers the clinical benefits of distinct aspects of clinical empathy – affective resonance, perspective-taking, compassion, imagining how – and presents a model of how these aspects work in an integrated way to yield the most effective empathic communication. I also address concerns about empathy and burnout

and argue for an “empathic curiosity” model of clinical empathy that may help practitioners avoid some of the risks of burnout that have been attributed to “sympathetic distress.”

The Evolution of Empathy Research

Increasing conceptual precision and improving empirical research on empathy are recent occurrences. A review of hundreds of earlier studies of empathy in healthcare reveals many different uses of the term “empathy” and very little precision in doing so (Pedersen, 2009). In addition, most of the empirical studies, unfortunately, asked doctors to self-report the degree to which they perceive themselves as generally empathic or even as having listened to a patient empathically – not necessarily a valid representation of their actually having *engaged* empathically.

In addition, even the higher quality research on empathy (psychology research outside of the medical setting) has tended to dichotomize affective and cognitive empathy by providing strictly affective or strictly informational stimuli (Zaki, Bolger, Ochsner, 2008). Yet patients undeniably bring both emotions and thoughts to their encounters with physicians, and for reasons I present below, it is crucial that they be met with empathy that integrates affective and cognitive elements.

Fortunately, over the past couple of decades the quality of empirical research on clinical empathy has improved, with more precise hypotheses and with improved methods that include direct observational studies in which doctor–patient interactions are videotaped over a year of clinical encounters (so that both parties adapt to videotaping in the clinic), as well as MRI and other studies (Decety, Smith, Norman, & Halpern, 2014; Finset, 2011; Suchman, Markakis, Beckman, & Frankel, 2011).

This is an exciting time to return to foundational questions about how empathy is therapeutic because we need to develop more accurate models of how emotional and cognitive aspects of empathy work together to create the dynamic, interpersonal communication that is essential for effective healthcare. That is, we need to get beyond academic questions about what goes on inside the mind of the empathizer in order to take a patient-centered perspective and address

how emotional and cognitive features enable patients to feel understood, develop trust, and experience greater agency – all of which are demonstrably therapeutic.

The Individual Aspects of Clinical Empathy and their Therapeutic Contributions

To begin investigating how the affective and cognitive aspects of empathy are related in the complex process of empathic engagement in a health care setting, it is important to delineate the four different aspects of clinical empathy. Their division into separate elements is, however, artificial, for in fact, they are usually (either wholly or partly) integrated. Some of them are interdependent, and all of them can enhance the others – hence “aspects,” implying elements of a larger, unified phenomenon.

The first aspect is *affective resonance*, which is also related to the more general phenomena of non-verbal attunement. Resonance is the familiar experience of a listener feeling an emotion, often at a reduced or subtle level, that corresponds to that of the speaker. This type of empathy has been shown, empirically, to play an important role in effective healthcare. For example, it can be crucial for enhancing history-taking and thus also diagnosis. Replicated observational studies of patient–physician interactions have shown that before patients talk about aspects of their history that are emotional, they give hints, often through gestures (Suchman et al., 1997). Clinicians who are non-verbally attuned respond to those hints, and then patients communicate (Finset, 2011). As an example, consider a woman oncologist whose young patient seems anxious not only about her own health after breast cancer surgery but also about something else. The oncologist feels a low level of “contagious” worry and communicates this to the patient non-verbally with raised eyebrows, expressing that she is not only sorry for what the patient is going through but is resonating with the patient’s feeling of anxiety. The patient discloses that she is worried about how all this will affect her marriage.

The second aspect of clinical empathy is what psychologists commonly call *perspective-taking*. This is seen in psychology research as a basic cognitive capacity to perceive a situation from another person’s point of view. There has been little research on perspective-taking in healthcare,



Professor Jodi Halpern

but it would seem to play a crucial role in helping patients with decision-making as well as other crucial aspects of healthcare. Research indicates that a lack of perspective-taking is one of the main triggers causing patients to file malpractice claims (Virshup, Oppenberg, & Coleman, 1999). In fields outside of healthcare, perspective-taking has been shown to play a core role in conflict resolution (Goldstein, Vezich, & Shapiro, 2014). Ethics and psychiatry consultants, who listen to countless cases in which patients are unhappy with medical care, understand that helping physicians consider the patient’s perspective is crucial for problem solving.

The claim that perspective-taking can reduce problematic breakdowns in patient–physician communication must acknowledge an inherent challenge, however: it is harder to take on another person’s perspective during a conflict, so getting to the beneficial effects of perspective-taking requires the additional skill of being able to stay interested in a point of view in conflict with your own. A related complication is that it can also be harder for physicians to perspective-take if resonance devolves into sympathetic distress. For example, medical students who become more personally distressed in response to patients’ distress have been shown to have steeper declines in cognitive empathy as their training progresses (Neumann et al., 2011). Even clinicians who stay emotionally engaged can sometimes find themselves feeling sympathetic distress, leading to outcomes that are less therapeutic.

In the example above, what role does perspective-taking play in empathic communication between the oncologist and the

young woman with breast cancer? Because the young woman's breast cancer is at a very early stage, surgery alone offers a greater than 95% chance of avoiding recurrence, so the oncologist believes the patient need not go through noxious radiotherapy. But the patient disagrees, and the oncologist accepts that for this particular patient even a 5% chance of recurrence is unacceptable. By being able to understand her patient's point of view, the oncologist will be able to help her make difficult decisions.

The third aspect of clinical empathy is *caring*, or compassion, feeling a mixture of good will towards and a desire to offer help (as opposed to pity) to the patient. Clinicians' express a range of benevolent emotions from concern or even worry to mindful compassion. There is increasing evidence that emotionally engaged clinicians – who patients perceive are genuinely worried about them when the situation warrants it– have greater therapeutic efficacy (Roter et al., 1998). This perceived emotional concern engenders trust, which is the most important predictor of adherence to treatment. Since about half of medical regimens are not followed, causing poor outcomes, improving adherence greatly improves the effectiveness of medical care. The empathic oncologist in our example conveys both her compassion – as a woman and fellow human she feels with the patient for having to face a serious health problem so young – and she feels and conveys a desire to help her and not abandon her.

The final process in the exercise of clinical empathy is *imagining how* an experience feels when one is inside it. This is the art of “imagining how...” rather than “knowing that.” My own work draws from the psychodynamic aspects of medical practice and from palliative care to emphasize that empathy involves the act of imagining what is significant from another person's perspective (Halpern, 2001). Imagining how is guided by curiosity to know what it feels like to be inside the patient's situation, and makes use of affective resonance to guide what the listener imagines. Imagining how someone else feels is like a daydream or fantasy, insofar as it has not only specific details, but also a connecting theme or mood that conveys that this is a happy, sad, exciting, frightening or other emotional experience. If a friend tells you that her spouse gave her a vacuum cleaner for Valentine's Day,

you will immediately imagine a very different scenario if you pick up on enthusiastic feelings or if you resonate with hurt feelings. In the clinical setting, resonance also helps patients give more information, as noted above, making it more likely for the listener to be able to fill in a more vivid picture of what the patient's concerns are really about.

How does curiosity and trying to imagine the experience of the patient help therapeutically? Insofar as the practitioner can better imagine what is salient for the patient, from that patient's perspective, she is more likely to get information that is crucial for determining effective and appropriate treatment. Second, it can help with diagnosis. (Why can't this patient get out of bed all day? Is the apparent lethargy depression or hypothyroidism or a cancer? What comes first, fatigue or demoralization?) Third, both effective medical care and appropriately helping patients with important decisions depend upon it. (Why doesn't this person with chronic depression take antidepressants? Why is this patient seeking surgery for back pain despite poor odds that it will help? Why does this person seek aid in dying?) Vividly imagining a patient's emotional experience goes beyond cognitive perspective-taking insofar as it is guided by resonance and thus feels like something affective is happening to the listener—she is actually daydreaming the patient's world and not just knowing facts about it.

How do practitioners take in enough of the details of the patient's world to be able to construct at least a momentary experiential grasp, or daydream-like experience? First, they have to be good listeners. Second, they have to develop the trust that allows patients to tell them more. We have already seen that resonance helps develop trust. So resonance enables patients to give more details, which then enables the doctor to picture things more fully, which she then can convey back to the patient, which in turn develops more trust. This is a virtuous cycle. The details gathered may not be particularly accurate at first, but as long as the practitioner lets the patient see that she is trying to understand things in detail rather than brush over them, and if she can invite the patient to point out mistakes and make corrections, the best empathic communication can happen.

To return to the example of the oncologist: she knows that her highly educated patient understands the medical facts, including how noxious the chemotherapy will be and that it introduces long term health risks. She is curious about how this patient's palpable fear may be contributing to her determination to pursue the most aggressive treatment. Based on their communication, she is able to imagine being inside the young woman's experience and comes to understand that the patient feels as if she's been cursed. She believes that having been unlucky enough to get cancer so young she will be unlucky again and end up in the 5% recurrence group. The oncologist helps the patient become conscious that this feeling of being singled out by fate is a common manifestation of fear. In grasping this, the patient begins to cry over the whole situation and eventually realizes that she cannot fully control the uncertainties in her future. She reconsiders her initial treatment decision.

Empathic Curiosity Enhances Therapeutic Benefits

The empathic curiosity model that I propose for the practice of clinical empathy involves all four of the aspects that I have outlined. It puts resonance to work in the service of learning more in order to better imagine how the patient is feeling and to communicatively connect in real time while co-imagining, in order to convey caring and appropriate worry.

Why is this model distinct from and of much greater therapeutic value than detached perspective-taking in clinical care?

First, being approached with detached curiosity makes people who are hearing bad news or are otherwise suffering feel like they are disappearing (Brison, 1996; Girgis & Sanson-Fisher, 1998; Ptacek, Fries, Eberhardt, & Ptacek, 1999). In contrast, empathic curiosity builds trust and is empowering (Halpern, 2001; Roter et al., 1998). Second, the model, as a whole, has a positive "cyclical" effect: affective engagement builds trust / resonance guides what the listener imagines / attuning of mood enables more relevant imagining how / imagining how improves communication in a virtuous cycle. In our case example, the patient senses the oncologist's affective resonance and that builds

trust; and for her part, the oncologist's resonance with the patient's anxiety leads her to imagine an anxious world view rather than a sad or a catastrophic one. All of this helps them better align their points of view and thus communicate more effectively.

Empathic curiosity and compassion supplement each other in empathic therapeutic encounters. On the one hand, curiosity without compassion or a pro-social attitude towards the patient could be used for dangerous or destructive purposes. On the other hand, generic compassion for all fellow humans lacks the curiosity about what, specifically, this patient is concerned about. These two situations represent degraded forms of empathic curiosity and compassion. In full empathic communication, on the other hand, the doctor's empathic curiosity about the individual patient is guided by a sense of shared humanity and shared possibilities for vulnerability and finitude.

Given that accepting one's vulnerability and finitude is not easy, the newer emphasis on training doctors in "mindful compassion" is a good step in the direction of facilitating empathy. Some hospitals, however, now translate "compassion" into "be kind," which devalues it and risks a return of paternalistic projections of "kindness," as exemplified by the phrase too often uttered by uncurious doctors: "I know how you feel." Empathic curiosity, in contrast, respects the individuality of each patient and says instead: "Tell me what I'm missing." Even a woman doctor who has had her own breast cancer and who thinks she knows how to be kind to a woman patient with breast cancer is at risk of missing important particulars of the patient's world.

One might wonder: What about avoiding burnout? In my view, the best path is to cultivate both empathic curiosity and mindful compassion, as the two work synergistically, as described above. We have enough research to take seriously the idea that compassion practices can help avoid burnout. We need more precise research on empathic curiosity, but we do have suggestive correlational findings that curiosity about another person's feelings helps caregivers reduce self-related anxiety (Decety et al., 2014). And all of us are familiar with the fact that becoming engaged with and curious about another person's

experience is the surest way to decenter from our own anxiety – consider how people relax by reading or watching television or film narratives of others’ lives. My decades of clinical observations and educating medical and nursing students have shown that those who were able to sustain interest in their practices were the ones who stayed curious about what their own emotional responses to difficult clinical encounters might be telling them about their patients. Of course this requires not only curiosity but also self-awareness, which is where the mindfulness component can be very useful. That developing curiosity and one’s empathic imagination reduces personal distress is reflected in a study that found that oncologists and hospice physicians with “exquisite empathy” showed the lower levels of burnout and distress than other physicians (Kearney, Weininger, Vachon, Harrison, & Mount, 2009).

For physicians to be mindful when they themselves are anxious helps the situation not devolve into over-identification, projection, or feeling helpless. On the other hand, when a doctor does feel helpless or merged, it is much better if she can recognize it and get help. Physicians who are unaware of their upset feelings are at risk of creating problems through everything from precipitous discharges to over or under treatment (Halpern, 2001).

In summary, the empathic curiosity model incorporates aspects of compassion in that it integrates a pro-social concern or feeling for the well-being of another with ongoing imagination work that attempts to grasp what the other person’s individual, subjective experience feels like from the inside out.

In Conclusion

Given this model, and the demonstrable effectiveness of its components, the question that follows is how to create an organizational culture that motivates empathic curiosity and emotional engagement so that patients can receive empathic care? Recent work by Jamil Zaki and others suggests how social expectations and environmental cues can readily motivate empathy or the lack of it (see Zaki, 2014). One important and underexplored area of social motivation is the expectations of patients who have been educated and empowered collectively to expect empathy.

An interdisciplinary group of us recently argued for the importance of expanding our research agenda beyond the longstanding focus on the intra-psychic aspects of empathy to include the dynamic, interpersonal aspects of successful empathic communication (Main, Walle, Kho, & Halpern, 2017). Achieving this would suggest the need to develop intervention studies to see if, in addition to educating physicians, empowering patients to expect and skillfully trigger empathy improves patient–physician communication. In this regard, we also need to work to eliminate systemic barriers to empathy, which include not only racism, sexism, gender and sexual orientation biases, and able-ism but also the specific devaluing of people with obesity, addictions or just “unhealthy” habits.

Over a decade ago I made a philosophical argument that when affective and cognitive empathy are integrated, the results are especially therapeutic (Halpern, 2001), and the empirical research that has been done since supports this. I hope this essay can help to bring more precision to the field of clinical empathy by uniting empirical findings and conceptual arguments in order to delineate testable hypotheses regarding which specific aspects of empathy are therapeutic. I also hope it helps to call into question the current shift in healthcare toward “feeling for” patients (compassion) and away from “feeling into” them (empathy). The rise of mindfulness and compassion need not be accompanied by the devaluing of empathic curiosity – in fact I hope I’ve shown that best practices integrate these approaches. Further, I continue to build on my hypothesis that the most powerful therapeutic/mutative factor in all the aspects of clinical empathy is rooted not in another person having compassion for you but in another person knowing how it feels to be in your world from the inside out. I hope to see both more precise research testing this hypothesis and the development of interventions to help physicians become better at empathic communication and imagining how it feels to be in their patient’s shoes.

References

- Brison, S. J. (1996). Outliving oneself: Trauma, memory and personal identity. In D. Meyers (Ed.), *Feminists Rethink the Self* (pp. 13-45).

- Boulder, CO: Westview Press.
- Decety, J., Smith, K., Norman, G., & Halpern, J. (2014). Clinical empathy: What can we learn from social and affective neuroscience? *World Psychiatry, 13*, 233-237.
- Finset, A. (2011). Research on person-centered clinical care. *Journal of Evaluation in Clinical Practice, 17*, 384-386.
- Girgis A., & Sanson-Fisher, R. W. (1998). Breaking bad news 1: Current advice for clinicians. *Behavioral Medicine, 24*, 53-59.
- Goldstein, N. J., Vezich, I. S., & Shapiro, J. R. (2014). Perspective-taking: When others walk in our shoes. *Journal of Personality and Social Psychology, 106*, 941-960.
- Halpern J. (2001). *From Detached Concern to Empathy: Humanizing Medical Practice*. Oxford, UK: Oxford University Press.
- Kearney, M. K., Weininger, R. B., Vachon, M.S., Harrison, R. L., Mount, B. M. (2009). Self-care of physicians caring for patients at the end of life: 'Being connected . . . A key to my survival. *Journal of the American Medical Association, 301*, 1155-1164.
- Main, A., Walle, E. A., Kho, C., Halpern, J (2017). The interpersonal functions of empathy: A relational perspective. *Emotion Review*.
- Neumann, M., Edelhäuser, F., Tauschel, D., Fisher, M. R., Wirtz, M., Woopen, C., Haramati, A., & Scheffer, C. (2011). Empathy decline and its reasons: A systematic review of studies with medical students and residents. *Academic Medicine, 86*, 996-1009.
- Pedersen, R. (2009). Empirical research on empathy in medicine – A critical review. *Patient Education and Counseling, 76*, 307-322.
- Ptacek, J. T., Fries, E. A., Eberhardt, T. L., Ptacek, J. J. (1999). Breaking bad news to patients: Physicians' perceptions of the process. *Support Care Cancer, 7*, 113-120
- Roter, D. L., Hall, J. A., Merisca, R., Nordstrom, B., Cretin, D., & Svarstad, B. (1998). Effectiveness of interventions to improve patient compliance: A meta-analysis. *Medical Care, 36*, 1138-1161.
- Suchman, A. L., Markakis, K., Beckman, H. B., Frankel, R. (1997). A model of empathic communication in the medical interview. *Journal of the American Medical Association, 277*, 678-682.
- Virshup, B. B., Oppenberg, A. A., Coleman, M. M. (1999). Strategic risk management: Reducing malpractice claims through more effective patient-doctor communication. *American Journal of Medical Quality, 14*, 153-159.
- Zaki, J. (2014). Empathy: A motivated account. *Psychological Bulletin, 140*, 1608–1647.
- Zaki, J., Bolger, N., Ochsner, K. (2008). It takes two: The interpersonal nature of empathic accuracy. *Psychological Science, 19*, 399-404.